

(3) STIC-Biotech/ChemLib

145029  
119

From: Li, Bao-Qun  
Sent: Monday, February 14, 2005 8:43 AM  
To: STIC-Biotech/ChemLib

PLEASE DO THE SEQUENCE HOMOLOGY AND INTERFERENCE OF SEQ ID NO; 20 IN APPLICATION SN. 09,664,363. THANKS.

Bao Qun Li M.D  
TC 1600  
Art Unit 1648  
Tel. 517-272-0904  
REM, 3C18  
Rm. 3D24

RECEIVED  
FEB 14 2005  
STIC-BIOTECH/CHEM LIBRARY  
(STIC)

\*\*\*\*\*

STAFF USE ONLY

Searcher: Arnold  
Searcher Phone: 2-2532  
Date Searcher Picked up: 2/12/05  
Date Completed: 2/12/05  
Searcher Prep/Rev. Time: 21:28:05  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA Sequence: # 1  
AA Sequence :# \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

*This Page Blank (upper)*



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 145029

**TO:** Bao-Qun Li  
**Location:** rem/3a24/3c18  
**Art Unit:** 1648  
**Monday, February 28, 2005**

**Case Serial Number:** 09/664363

**From:** Deirdre Arnold  
**Location:** Biotech-Chem Library  
REM 1A64  
**Phone:** 571-272-2532

**[Deirdre.Arnold@uspto.gov](mailto:Deirdre.Arnold@uspto.gov)**

### Search Notes

*Please feel free to contact me if you have any questions or would like to amend the search.*

Thank you for using STIC services.

Regards,  
Deirdre Arnold

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 02:17:34 ; Search time 8554 Seconds

11572.832 Million cell updates/sec

Title: US-09-664-363-20

Perfect score: 2043

Sequence: 1 TCGGAGGGCGCTTCAGG.....TGTGGCCATGAGATGGC 2043

Scoring table: IDENTITY\_NUC Gapop 10.0 , Gapext: 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl,\*

- 1: gb\_ba:\*
- 2: gb\_ng:\*
- 3: gb\_in:\*
- 4: gb\_on:\*
- 5: gb\_ex:\*
- 6: gb\_pat:\*
- 7: gb\_ph:\*
- 8: gb\_pl:\*
- 9: gb\_pr:\*
- 10: gb\_ro:\*
- 11: gb\_sts:\*
- 12: gb\_sy:\*
- 13: gb\_un:\*
- 14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2043	100.0	A28153	A28153 PT-NANBH
2	2043	100.0	A32201	A32201 NANBH Pt_p
3	2043	100.0	A28149	A28149 Sequence
4	2043	100.0	A28157	A28157 PT-NANBH MR
5	2043	100.0	A28151	A28151 Sequence
6	2035	99.6	A28150	A28150 Sequence
7	2022.2	99.0	A32203	A32203 NANBH Pt_p
8	1804.6	88.3	A28153	A28153 PT-NANBH
9	1802.6	88.2	A28157	A28157 PT-NANBH
10	1801.4	88.2	A28154	A28154 Sequence
11	1798.2	88.0	A28155	A28155 Sequence
12	1796.6	87.9	A28156	A28156 Sequence
13	1796.6	87.9	A28157	A28157 Sequence
14	1796.6	87.9	A28158	A28158 Sequence
15	1796.6	87.9	A28159	A28159 Sequence
16	1796.6	87.9	A28160	A28160 Sequence
17	1796.6	87.9	A28161	A28161 Sequence
18	1796.6	87.9	A28162	A28162 Sequence
19	1796.6	87.9	A28163	A28163 Sequence
20	1796.6	87.9	A28164	A28164 Sequence
21	1796.6	87.9	A28165	A28165 Sequence
22	1796.6	87.9	A28166	A28166 Sequence
23	1796.6	87.9	A28167	A28167 Sequence
24	1796.6	87.9	A28168	A28168 Sequence
25	1796.6	87.9	A28169	A28169 Sequence
26	1796.6	87.9	A28170	A28170 Sequence
27	1796.6	87.9	A28171	A28171 Sequence
28	1796.6	87.9	A28172	A28172 Sequence
29	1796.6	87.9	A28173	A28173 Sequence
30	1795	87.9	A28174	A28174 Sequence
31	1795	87.9	A28175	A28175 Sequence
32	1795	87.9	A28176	A28176 Sequence
33	1795	87.9	A28177	A28177 Sequence
34	1793.4	87.8	A28178	A28178 Sequence
35	1793.4	87.8	A28179	A28179 Sequence
36	1793.4	87.8	A28180	A28180 Sequence
37	1788.6	87.5	A28181	A28181 Sequence
38	1788.6	87.5	A28182	A28182 Sequence
39	1787	87.5	A28183	A28183 Sequence
40	1787	87.5	A28184	A28184 Sequence
41	1787	87.5	A28185	A28185 Sequence
42	1787	87.5	A28186	A28186 Sequence
43	1787	87.5	A28187	A28187 Sequence
44	1787	87.5	A28188	A28188 Sequence
45	1787	87.5	A28189	A28189 Sequence

#### ALIGNMENTS

RESULT 1	A28153	2043 bp	DNA linear	PAT 07-JUN-1995
LOCUS	A28153	PT-NANBH	mRNA Fragment from patent	GB2239245.
DEFINITION				
ACCESSION	A28153			
VERSION	A28153.1	GL:1249636		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
POST-TRANSFUSIONAL NON-A NON-B HEPATITIS VIRAL POLYPEPTIDES				
PATENT: GB 2239245-A 20 26-JUN-1991;				
1. . 2043				
/organism="synthetic construct"				
/mol_type="unassigned DNA"				
/db_xref="taxon:32630"				
/db_xref="clone:156/192"				
<1 .. >2043				
/codon_start=1				
/transl_table=11				
/product="PT-NANBH polypeptide"				
/protein_id=CAA01930.1"				
/db_xref="GI:1249637"				
/translation="MEGWETLGLPHDAFLSQTQKAGDNFPVLYAVOTVACRAQAPP				
PSWQMKCKLIRKPLTPIKPLTPIKPLTPIKPLTPIKPLTPIKPLTPIKPLTPIKPLT				
WLVUGGVLAALAYCLTGSSVWVYGRILIGRAIVPVERVILQFVNLQVYQFDEMEBCASHLP				
IEQCMQLAQPKOKAQLGILQATQKAEAAAPVYESKORALETEWAKHWNNTFSGIQL				
AGUSTLPGNPAASLMATTAATSVSPLTQSTLTLNLLGGWAQLOPPSAVNGVAG				
TAGAVGVSIGLGKVTLILQAGVAGVAGLVAKVMSCSEMPSDIDUMLPAILSPGA				
APF33324 Hepatitis				
AF16560 Hepatitis				
BLTTQQLRERLHQWNPDCGSTSGSWRDWMDICLTFKWLQSKLRLPQV				
FFSCQRYKGWVQWGDGMQTTSCGQAQTGKRSIVGPKTCNSNNWHGTPINP				
TRPCTPSAPNYSRALVRAAEYVEVTRVDPHYTMSTMDVNUVKCPQVAPERP				
EVGVVLRHRKAPACKLILREBFVTFQVGNQVLYQSQPCPERDVAVLTSMDPSH				
TAETAKRRLARGBPSLASSASQSLSDPSSKATYTQNDPFDADLIEANLWRHEMG"				
ORIGIN				
Query Match	100.0%	Score 2043;	DB 6;	Length 2043;
Best Local Similarity	100.0%	Pred. No. 0;		
Result No.	1			
Score	2043			
Match Length	100.0			
DB ID	A28153			
Description	PT-NANBH			

Matches	2043;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	TGGAGGGCGCTTCAGGCGTACCCACCTGGATGCCACTTCTGTCCAAACAAAG							
Db	1	TTGGAGGGCGCTTCAGGCGTACCCACCTGGATGCCACTTCTGTCCAAACAAAG							
Qy	61	CAGGAGAGAGAACACTTCCCCTACCTGGTGGCTACGGGTTACGGGTTACGGG							
Db	61	CAGGAGAGAGAACACTTCCCCTACCTGGTGGCTACGGGTTACGGGTTACGGG							
Qy	121	CAGGCCCAACTCCCATGGATCAAATGTTGAAGGTCTGAGGTACTCTGGCTAGGGC							
Db	121	CAGGCCCAACTCCCATGGATCAAATGTTGAAGGTCTGAGGTACTCTGGCTAGGGC							
Qy	121	ACACACCCATACCCATGGATCAAATGTTGAAGGTCTGAGGTACTCTGGCTAGGGC							
Db	121	ACACACCCATACCCATGGATCAAATGTTGAAGGTCTGAGGTACTCTGGCTAGGGC							
Qy	181	CTGGCGGGGCACACCCCTGGTGTAGGTGGGACCCGTCACAAACAGGGTACCCCT							
Db	181	CTGGCGGGGCACACCCCTGGTGTAGGTGGGACCCGTCACAAACAGGGTACCCCT							
Qy	241	AGCACCTGGTGTGGGGGGCTTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG							
Db	241	AGCACCTGGTGTGGGGGGCTTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG							
Qy	301	AGCACCTGGTGTGGGGGGCTTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG							
Db	301	AGCACCTGGTGTGGGGGGCTTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG							
Qy	361	GGCAGGGCATTTGGGGGGACATCTGCTGGGGGGCTATGTTCCGAC							
Db	361	GGCAGGGCATTTGGGGGGACATCTGCTGGGGGGCTATGTTCCGAC							
Qy	421	AGGAAAGTCTTACCGAGGTCTGATGAGTGGAGAGTGGAGAGTGGAGACTCC							
Db	421	AGGAAAGTCTTACCGAGGTCTGATGAGTGGAGAGTGGAGAGTGGAGACTCC							
Qy	481	ATGACGAGGAGATGAGCTGGAGAGTCAACAAAGGCTCGCTGCGAG							
Db	481	ATGAGGAGGAGATGAGCTGGAGAGTCAACAAAGGCTCGCTGCGAG							
Qy	541	ACAGCCCCAGCAACGGGAGCCGGCTCGTGGGTCAAGTGGAGSCCTT							
Db	541	ACAGCCCCAGCAACGGGAGCCGGCTCGTGGGTCAAGTGGAGSCCTT							
Qy	601	GAGACCTCTGGGAGAACATGTGAACTCATGGGATAAGTAGTGTAGAGC							
Db	601	GAGACCTCTGGGAGAACATGTGAACTCATGGGATAAGTAGTGTAGAGC							
Qy	661	TGTCGACTCTGGGAGAACATGTGAACTCATGGGATAAGTAGTGTAGAGC							
Db	661	TGTCGACTCTGGGAGAACATGTGAACTCATGGGATAAGTAGTGTAGAGC							
Qy	721	ACTAGCGCGTCACTCCCAACTACCTCTGGTAACTACCTGGGGATGGTAGCC							
Db	721	ACTAGCGCGTCACTCCCAACTACCTCTGGTAACTACCTGGGGATGGTAGCC							
Qy	781	GCCCAACTCTGGGAGAACATGTGAACTCATGGGATAAGTAGTGTAGAGC							
Db	781	GCCCAACTCTGGGAGAACATGTGAACTCATGGGATAAGTAGTGTAGAGC							
Qy	841	GCTGTTGGCAGATAAGGCTTGGGAGGTGTGAGCTGGAGACATTTGGGAGA							
Db	841	GCTGTTGGCAGATAAGGCTTGGGAGGTGTGAGCTGGAGACATTTGGGAGA							
Qy	901	GGAGTGCGAGGGCGCTGGCTGGAGCTGGAGACATTTGGGAGA							
Db	901	GGAGTGCGAGGGCGCTGGAGCTGGAGACATTTGGGAGA							
Qy	961	GACCTGGTAACTACTCCCGGCTAGGGAGGGGGTGGGGTGGTG							
Db	961	GACCTGGTAACTACTCCCGGCTAGGGAGGGGGTGGGGTGGTG							
Qy	1021	TGGCAGGGAGACTGGCGACGGGGGGCTGGAGCTGGAGCTGGAGAC							
Db	1021	TGGCAGGGAGACTGGCGACGGGGGGCTGGAGCTGGAGAC							
RESULT	2	A32201	2043 bp	DNA	linear	PAT	10-DEC-1996	LOCUS	



QY	1621	GATTTTCACTACTGTCAGACGATGACCACAGAACGTAATTGGCGTGCAAGGTTCA	1680	Db
Db	1621	GATTTCACATGAGTCAGCAGCATGACACTGACAAGTAATGCGCGTCCAGGTCA	1680	QY
QY	1681	GCCCCCGAAATCTTCACAGAGTGGATGGGGCGCTCGCACAG3TAACGCTCGGGTC	1740	Db
Db	1681	GCCCCGAATCTTCACAGAGTGGATGGGGCGCTCGCACAG3TAACGCTCGGGTC	1740	QY
QY	1741	AAACCTCTCTACAGGGAGGTCACTCCAGGGCTCAACCAATACTGTTGGG	1800	Db
Db	1741	AAACCTCTCTACAGGGAGGTCACTCCAGGGCTCAACCAATACTGTTGGG	1800	QY
QY	1801	TGGGCTCCATGAGGGGAAACGGATAGCGATGCTGCTCACTTCATCTCCGC	1860	Db
Db	1801	TGGGCTCCATGAGGGGAAACGGATAGCGATGCTGCTCACTTCATCTCCGC	1860	QY
QY	1861	CCCTCCCACATCACGGCAGAGGCTAAGGCCAGCTGGCAGGATGAGTGTG	1920	Db
Db	1861	CCCTCCCACATCACGGCAGAGGCTAAGGCCAGCTGGCAGGAGTGTG	1920	QY
QY	1921	TGGGCTCCATGAGGGGAAACGGATAGCGATGCTGCTCACTTCATCTCCGC	1980	Db
Db	1921	TGGGCTCCATGAGGGGAAACGGATAGCGATGCTGCTCACTTCATCTCCGC	1980	QY
QY	1981	CAAAATGACTCCGAGGCTGACTTCATGGCCAACTCTGGGGATGAGAT	2040	Db
Db	1981	CAAAATGACTCCGAGGCTGACTTCATGGCCAACTCTGGGGATGAGAT	2040	QY
QY	2041	GGC 2043	2041	Db
QY	2041	GGC 2043	2041	Db
RESULT 3				
AR144049	AR144049	Sequence 20 from patent US 6210675.	2043 bp	DNA
DEFINITION	Sequence 20 from patent	US 6210675.		linear
VERSION	AR144049			PAT 08-AUG-2001
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 2043)			
AUTHORS	Highfield, P. Edmund., Rodgers, B. Colin., Tedder, R. Seton. and			
TITLE	PT-NANB hepatitis polypeptides			
JOURNAL	Patent: US 621065-A 2003 APR-2001;			
FEATURES	Location/Qualifiers			
source	1..2043 /organism="unknown" /mol_type="unassigned DNA"			
ORIGIN				
Query Match	100.0%	Score 2043; DB 6; Length 2043;		
Best local Similarity	100.0%	Score 2043; DB 6; Length 2043;		
Matches 2043; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;		
QY	1	TGGGGGGGCTTCACAGGCTCACCCACCTGGTGGGACAGGTACTCTGGTGTGCAACAAAG	60	Db
Db	1	TGGGGGGGCTTCACAGGCTCACCCACCTGGTGGGACAGGTACTCTGGTGTGCAACAAAG	60	QY
QY	61	CAGGGAGGACACCTCCCTACTGGTGGGACAGGTACTCTGGTGTGCACTAGGGC	120	Db
Db	61	CAGGGAGGACACCTCCCTACTGGTGGGACAGGTACTCTGGTGTGCACTAGGGC	120	QY
QY	121	CAGGGCCACCTCCATGGGATCAAATGTTGAGGTCTATCGCTAAAGGCTACT	180	Db
Db	121	CAGGGCCACCTCCATGGGATCAAATGTTGAGGTCTATCGCTAAAGGCTACT	180	QY
QY	181	CTGGCGGGCAACACCTGGTGTAGGTGGAGGCGTCAAACGAGGTCACTTC	240	Db
QY	241	ACACACCCATAACAAATCATGCGCAGTCGTCGCCGACTCTGGGGTC	300	Db
Db	241	ACACACCCATAACAAATCATGCGCAGTCGTCGCCGACTCTGGGGTC	300	QY
QY	301	AGCACCTGGCTGGTGGGGCTCTGCAACTCTGGCTGCTGATATGCTGACAA	360	Db
Db	301	AGCACCTGGCTGGTGGGGCTCTGCAACTCTGGCTGCTGATATGCTGACAA	360	QY
QY	360	AGCACCTGGCTGGTGGGGCTCTGCAACTCTGGCTGCTGATATGCTGACAA	360	Db
Db	360	AGCACCTGGCTGGTGGGGCTCTGCAACTCTGGCTGCTGATATGCTGACAA	360	QY
QY	420	GGCAGCTGGCTATGTTGGGAGGTGATCATCTTGGCGGGGGCTATGTTG	480	Db
Db	420	GGCAGCTGGCTATGTTGGGAGGTGATCATCTTGGCGGGGGCTATGTTG	480	QY
QY	480	ATCGAGCAGGAGAATGCGAGCTGGCGAGGAGCTGAGCAAAAGGCTCGG	540	Db
Db	480	ATCGAGCAGGAGAATGCGAGCTGGCGAGGAGCTGAGCAAAAGGCTCGG	540	QY
QY	540	ACAGCCACCAAGGAGGGGGCTCTCCCTGGAGGAGCTGAGGAGCTGCG	600	Db
Db	540	ACAGCCACCAAGGAGGGGGCTCTCCCTGGAGGAGCTGAGGAGCTGCG	600	QY
QY	600	GAGACCTCTGGGAAACACATGTTGAACTCATCGGGGATGAGTCAGG	660	Db
Db	600	GAGACCTCTGGGAAACACATGTTGAACTCATCGGGGATGAGTCAGG	660	QY
QY	660	TGGGACTCTGGGGATCCGGGATCACTGAGGTTGAGTCAGGCTGTG	720	Db
Db	660	TGGGACTCTGGGGATCCGGGATCACTGAGGTTGAGTCAGGCTGTG	720	QY
QY	720	ACTAGCCCCCTCACCACCACTACCCCTCTGTAACTCCCTGGGGATG	780	Db
Db	720	ACTAGCCCCCTCACCACCACTACCCCTCTGTAACTCCCTGGGGATG	780	QY
QY	780	GCCTAACCTGCTCCCCCTGGCTGCTGCTGCTGCTGCTGCTG	840	Db
Db	780	GCCTAACCTGCTCCCCCTGGCTGCTGCTGCTGCTGCTGCTG	840	QY
QY	840	GCCTGGGAGCATGGCTGGGGCTGAGGCTGCTGAGGAGCTATGAG	900	Db
Db	840	GCCTGGGAGCATGGCTGGGGCTGAGGCTGCTGAGGAGCTATGAG	900	QY
QY	900	GGAGTGGAGGGGGCTCTGGCTTAAAGGTCTAGGGCGGAATGCC	960	Db
Db	900	GGAGTGGAGGGGGCTCTGGCTTAAAGGTCTAGGGCGGAATGCC	960	QY
QY	960	GACCTGGTTAACTTAACTCCGGCTCTGGCTTAAGGTCTAGGGCG	1020	Db
Db	960	GACCTGGTTAACTTAACTCCGGCTCTGGCTTAAGGTCTAGGGCG	1020	QY
QY	1020	TGGCGAGGCTACTCGTGGGAGGTGGGCTTCAAGGGGGCTGTCG	1080	Db
Db	1020	TGGCGAGGCTACTCGTGGGAGGTGGGCTTCAAGGGGGCTGTCG	1080	QY
QY	1080	CGGTGTAGTGGTGGCTGGGGGTAACGATGTTCCCTGGCTCC	1140	Db
Db	1080	CGGTGTAGTGGTGGCTGGGGGTAACGATGTTCCCTGGCTCC	1140	QY
QY	1140	AGCACGCCCCAGCTGGTGTGCTGACAGTCTCTGACCTTACTAT	1200	Db
Db	1140	AGCACGCCCCAGCTGGTGTGCTGACAGTCTCTGACCTTACTAT	1200	QY
QY	1200	AAGGGCTCACCAGTGGATAACGGAGCTGGCTCACGCCCTGTC	1260	Db
Db	1200	AAGGGCTCACCAGTGGATAACGGAGCTGGCTCACGCCCTGTC	1260	QY
QY	1260	AGGAGTGGTGGGACTGGAATGAGCTGACTCAAGACCTGCTC	1320	Db
Db	1260	AGGAGTGGTGGGACTGGAATGAGCTGACTCAAGACCTGCTC	1320	QY
QY	1320	AGGAGTGGTGGGACTGGAATGAGCTGACTCAAGACCTGCTC	1320	Db





Db	241	ACACACCCATACCAAAATTCTATCATGCCATGCAGCAGCCACCTGGAGGTGTCACG	300	Qy	1381	GCTGGCGGGAGACGCCATCATGCCAGCACACACTGCTCATGGAGCACAGATACGGGA	1440		
Qy	301	AGCACCTGGTGTCTGGTGGCGGGCTCTTGAGCTCTGCTGGCTGCTGCTGTTGGG	360	Db	1381	GTGCGGGAGACGCCATCATGCCAGCACACACTGCTCATGGAGCACAGATACGGGA	1440		
Db	301	AGCACCTGGTGTCTGGTGGCGGGCTCTTGAGCTCTGCTGGCTGCTGCTGTTGGG	360	Qy	1441	CATGTCAAAAGGTTCCATAGGGATGTTGGGCTTAAGCTGTGTTAATGTCGCT	1500		
Qy	361	GCGAGCTGTGCTTGGGGTAGGATCTCTTGTCTGGGGGGGGCTATGTTCCGGAC	420	Db	1441	CATGTCAAAAGGTTCCATAGGGATGTTGGGCTTAAGCTGTGTTAATGTCGCT	1500		
Db	361	GCGAGCTGTGCTTGGGGTAGGATCTCTTGTCTGGGGGGGGCTATGTTCCGGAC	420	Qy	1501	GGACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Qy	421	AGGGAAGTCCTTACCAAGGAGTGTGATGAGATGAGAAGAGGTGGGGTGCACCTCCCTAC	480	Db	1501	GGACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Db	421	AGGGAAGTCCTTACCAAGGAGTGTGATGAGATGAGAAGAGGTGGGGTGCACCTCCCTAC	480	Qy	1511	GERACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Qy	481	ATGAGGAGGAGATGAGCTCCGGAGGAGTGTGATGAGATGAGAAGAGGTGGGGTGCAG	540	Db	1511	GERACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Db	481	ATGAGGAGGAGATGAGCTCCGGAGGAGTGTGATGAGATGAGAAGAGGTGGGGTGCAG	540	Qy	1521	GATTCCACTACTGTGAGGAGTGACCAACTGAAATGCGGTSAGGTTCCA	1680		
Qy	541	ACAGCCGCCAACCAASCGGAAACCGGACCTCCGTGAGCTGGAGCTCCAGTGGAGCCT	600	Db	1521	GATTCCACTACTGTGAGGAGTGACCAACTGAAATGCGGTSAGGTTCCA	1680		
Db	541	ACAGCCGCCAACCAASCGGAAACCGGACCTCCGTGAGCTGGAGCTCCAGTGGAGCCT	600	Qy	1561	TATTCGAGGCCCTGAGCGGGTGGTGTGCTGAGGAGTACGGGTTACCGGGGGGG	1620		
Qy	601	GAGACCTTCTGGGAAACACAGTGTGAACTCATCGGGATACTGACTTAGEAGGC	660	Db	1561	TATTCGAGGCCCTGAGCGGGTGGTGTGCTGAGGAGTACGGGTTACCGGGGGGG	1620		
Db	601	GAGACCTTCTGGGAAACACAGTGTGAACTCATCGGGATACTGACTTAGEAGGC	660	Qy	1621	TATTCGAGGCCCTGAGCGGGTGGTGTGCTGAGGAGTACGGGTTACCGGGGGGG	1620		
Qy	661	TATTCGAGCTCCTGGGAACTCCGGGATTCGGGATCATCTGGGGATGGTAGCC	720	Db	1621	TATTCGAGGCCCTGAGCGGGTGGTGTGCTGAGGAGTACGGGTTACCGGGGGGG	1620		
Db	661	TATTCGAGCTCCTGGGAACTCCGGGATTCGGGATCATCTGGGGATGGTAGCC	720	Qy	1681	GCCCCCRAATTCTTCAAGAGTGGGCTGACGCTTCAGGTTAGGGTGGGGGGGG	1740		
Qy	721	ACTAGCCGCGTACCCACCAACTACCTCTCTCTACACCTCTGGGGATGGTAGCC	780	Db	1681	GCCCCCRAATTCTTCAAGAGTGGGCTGACGCTTCAGGTTAGGGTGGGGGGGG	1740		
Db	721	ACTAGCCGCGTACCCACCAACTACCTCTCTCTACACCTCTGGGGATGGTAGCC	780	Qy	1741	AACACTCTCTAAGGGGGAGGTCACTTCAGGTCGGGTCACAGCTCTCTCTCT	1800		
Qy	781	GCCCCAATCTGGTCCCTCCAGGAGTGTGCTCAAGTTCTGGGGGGGGGGGGGGGG	840	Db	1741	AACACTCTCTAAGGGGGAGGTCACTTCAGGTCGGGTCACAGCTCTCTCTCT	1800		
Db	781	GCCCCAATCTGGTCCCTCCAGGAGTGTGCTCAAGTTCTGGGGGGGGGGGGGGGG	840	Qy	1801	TGCACTCTCCATGCGGAGCCGAGGAGTGTGCTGTCACTTCAGTCACCGAC	1860		
Qy	841	GCTGTTGCGCAAGCATAGGGCTTCTGGGAGGTGTGCTGACATCTGGGGGGGG	900	Db	1801	TGCACTCTCCATGCGGAGCCGAGGAGTGTGCTGTCACTTCAGTCACCGAC	1860		
Db	841	GCTGTTGCGCAAGCATAGGGCTTCTGGGAGGTGTGCTGACATCTGGGGGGGG	900	Qy	1861	CCCTCCACACACAGAGAGAGGGTAAAGGGAGCTGGGGGGGGGGGGGGGGGG	1920		
Qy	901	GGAGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	960	Db	1861	CCCTCCACACACAGAGAGAGGGTAAAGGGAGCTGGGGGGGGGGGGGGGGGG	1920		
Db	901	GGAGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	960	Qy	1921	TGGCCAGCTCTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1980		
Qy	961	GACCTGGTTAACTACTCCCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020	Db	1921	TGGCCAGCTCTAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1980		
Db	961	GACCTGGTTAACTACTCCCTCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020	Qy	1981	CAAATTCATCTCCAGGGCTGACCTCTCAGGGCAACCTCTCTGGGGGGGGGGGG	2040		
Qy	1021	TGGCAGCGATRACTGCTCGGCACTGCTGGGGGGGGGGGGGGGGGGGGGGGG	1080	Db	1981	CAAATTCATCTCCAGGGCTGACCTCTCAGGGCAACCTCTCTGGGGGGGGGGGG	2040		
Db	1021	TGGCAGCGATRACTGCTCGGCACTGCTGGGGGGGGGGGGGGGGGGGGGGGG	1080	Qy	2041	GCC 2043			
Qy	1080	A76577 RESULT 6		Db	2041	GCC 2043			
Db	1080	A76577 LOCUS A76577		Qy	1381	GCTGGCGGGAGACGCCATCATGCCAGCACACACTGCTCATGGAGCACAGATACGGGA	1440		
Qy	1080	DEFINITION Sequence 9 from Patent WO9317110.		Db	1381	GCTGGCGGGAGACGCCATCATGCCAGCACACACTGCTCATGGAGCACAGATACGGGA	1440		
Db	1080	ACCESSION A76577		Qy	1441	CATGTCAAAAGGTTCCATAGGGATGTTGGGCTTAAGCTGTGTTAATGTCGCT	1500		
Qy	1081	VERSION A76577.1		Db	1441	CATGTCAAAAGGTTCCATAGGGATGTTGGGCTTAAGCTGTGTTAATGTCGCT	1500		
Db	1081	KEYWORDS unidentifed		Qy	1501	GGACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Qy	1140	SOURCE unidentifed		Db	1501	GGACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Db	1140	ORGANISM unclassified		Qy	1511	GERACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Qy	1141	REFERENCE 1 (bases 1 to 7065)		Db	1511	GERACATCCCCATCAAGCATACACACGGGGCTTGACGGTGTGACCCCTCCAGGCCAAC	1560		
Db	1141	AUTHORS Rodgers, B.C. and Parker, D.		Qy	1521	GATTCCACTACTGTGAGGAGTGACCAACTGAAATGCGGTSAGGTTCCA	1680		
Qy	1201	TITLE RBCOMBINANT HEPATITIS C VIRUS POLYPEPTIDE		Db	1521	GATTCCACTACTGTGAGGAGTGACCAACTGAAATGCGGTSAGGTTCCA	1680		
Db	1201	PARENT: WO 9317110-A 9 02-SEP-1993; WELLCOME FOUND (GB)		Qy	1561	TATTCGAGGCCCTGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1620		
Qy	1260	FEATURES source		Db	1561	TATTCGAGGCCCTGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1620		
Db	1260	1 . . . 7065 /organism="unidentified"		Qy	1621	TATTCGAGGCCCTGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1620		
Qy	1321	/mol type="unassigned DNA"		Db	1621	TATTCGAGGCCCTGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1620		
Db	1321	/db_xref="taxon:37644"		Qy	17065	<1. . . 7065 /codon_start=1 /note="Encodes putative PT-NAMB non-structural proteins"			
Qy	1380	/codon_start=1		Db	17065	<1. . . 7065 /codon_start=1			



Db 3604 GTCCTGGGGGAGACGCCATCATGCAGACCCCTGCTCATGGACACAGATCACCGA 3663  
 Qy 1441 CATGTCAAAMACGGTTCATAGGATGTTGGCCTTAAGACTGTGTAACATGTCAT 1500  
 Db 3664 CATGTCAAAMACGGTTCATAGGATGTTGGCCTTAAGACTGTGTAACATGTCAT 3723  
 Qy 1501 GGAACTTCCCATCAACGCAACACACAGGCGGCTGAGCCCTGGCGGCAAAC 1560  
 Db 3724 GGAACATTCCCATCAACGCAACACAGGCGGCTGAGCCCTGGCGGCAAAC 3783  
 Qy 1561 TATTCCAGGGGCTGTGGGGTGGCTGCTAGAGTAGCTGGAGGTACCGGGGG 1620  
 Db 3784 TATTCCAGGGGCTGTGGGGTGGCTGCTAGAGTAGCTGGAGGTACCGGGGG 3843  
 Qy 1621 GATTTCACTAGTGAGGAGATGACCTACAACTACAACGTTAACGCTTCA 1680  
 Db 3844 GATTTCACTAGTGAGGAGATGACCTACAACTACAACGTTAACGCTTCA 3903  
 Qy 1681 GCCCCCGAATTCTTCACAGAGTGAGTGGAGGGGGGGGGGGGGGGGGGG 1740  
 Db 3904 GCCCCCGAATTCTTCACAGAGTGAGTGGAGGGGGGGGGGGGGGGGGGG 3963  
 Qy 1741 AACCTCTCTACGGGAGGAGGTCACATCCAGGTGGGTCACCAAATCCTGGTGG 1800  
 Db 3964 AACCTCTCTACGGGAGGAGGTCACATCCAGGTGGGTCACCAAATCCTGGTGG 4023  
 Qy 1801 TCGCAGCTCCATGCGAGCCGAAACCGATGAGCTGCTCACTTCATGTCACCGAC 1860  
 Db 4024 TCGCAGCTCCATGCGAGCCGAAACCGATGAGCTGCTCACTTCATGTCACCGAC 4083  
 Qy 1861 CCCTCCACATCACAGAGGAGGGTAAGGCAGGCTGGCAGGGGTCTCCCTCC 1920  
 Db 4084 CCCTCCACATCACAGAGGAGGGTAAGGCAGGCTGGCAGGGGTCTCCCTCC 4143  
 Qy 1921 TTGGCAGCTTCAGTAGGAGCTGCTGGCCTCTGGCTCTGGAGGGACATACATTAC 1980  
 Db 4144 TTGGCAGCTTCAGTAGGAGCTGCTGGCCTCTGGCTCTGGAGGGACATACATTAC 4203  
 Qy 1981 CAAATGACTTCCAGAGGCTGACCTCATCAGGCCAACCTCTGTCGCGCATGAGAT 2040  
 Db 4204 CAAATGACTTCCAGAGGCTGACCTCATCAGGCCAACCTCTGTCGCGCATGAGAT 4263  
 Qy 2041 GGC 2043  
 Db 4264 GGC 4266

## ORIGIN

Query Match 99.0%; Score 2022.2; DB 6; Length 3750;

Best Local Similarity 99.4%; Pred. No. 0; Matches 2030; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Db 1 TGGAGGGCGCTTCAGGGCTCACCCAGCTGGATGCCCACCTCTGGTCCAAACAAAG 60  
 Db 1 TGGAGGGCGCTTCAGGGCTCACCCAGCTGGATGCCCACCTCTGGTCCAAACAAAG 60  
 Db 61 CAGGCAGGAGAACCTCCCTACTCTGCTGGCTGAGGCTACTCTGTCGCTGAGGCC 120  
 Db 61 CAGGCAGGAGAACCTCCCTACTCTGCTGGCTGAGGCTACTCTGTCGCTGAGGCC 120  
 Db 121 CAGGCCAACCTCCATGAGGCTAAATGAGGAGTGTCTCATGGCTAACGCTACT 180  
 Db 121 CAGGCCAACCTCCATGAGGCTAAATGAGGAGTGTCTCATGGCTAACGCTACT 180  
 Qy 181 CTGGCGGGCAACACCCCTCTGATAGGCTGGGAGCCGCTCCAAAAGAGGTACCC 240  
 Db 181 CTGGCGGGCAACACCCCTCTGATAGGCTGGGAGCCCTCAAAAAGAGGTACCC 240  
 Qy 241 ACACACCCATACCAAAATCATCATGCGCATCATCAGGCCACCTGGCTCGTACCG 300  
 Db 241 ACACACCCATACCAAAATCATCATGCGCATCATCAGGCCACCTGGCTCGTACCG 300  
 Qy 301 AGCACCTGGGCTGCTGGGGGGCTCTGCGAGCTCTGCTGCTGCTGCTGACACA 360  
 Db 301 AGCACCTGGGCTGCTGGGGGGCTCTGCGAGCTCTGCTGCTGCTGACACA 360  
 Qy 361 GGCAGCTGGTCATGGGAGGATCATCTGTCGGGGGGGGCTATGTCGGAC 420  
 Db 361 GGCAGCTGGTCATGGGAGGATCATCTGTCGGGGGGGGCTATGTCGGAC 420  
 Qy 421 AGGGAACTCTTACAGGGAGTCGATGAGATGGAGTCGACCTCCCTAC 480  
 Db 421 AGGGAACTCTTACAGGGAGTCGATGAGATGGAGTCGACCTCCCTAC 480  
 Qy 481 ATGCACGAGCATGAGCTGGCCAGAGCTTCAGGAAAGGCTCGGTTGAGCAG 540  
 Db 481 ATGCACGAGCATGAGCTGGCCAGAGCTTCAGGAAAGGCTCGGTTGAGCAG 540  
 Qy 541 ACAGCCCAAGCAAGGGGGCTGTCGGGGCTGCAAGCTGGGGATACAGTACTAGGGC 660  
 Db 541 ACAGCCCAAGCAAGGGGGCTGTCGGGGCTGCAAGCTGGGGATACAGTACTAGGGC 660  
 Qy 601 GAGACCTCTGGCGAAACATGTTGGGACTTCATCGGGGATACAGTACTAGGGC 660  
 Db 601 GAGACCTCTGGCGAAACATGTTGGGACTTCATCGGGGATACAGTACTAGGGC 660  
 Qy 661 TTGTCCTCTGGCGCTGGAGTCCCGGCTGTCAGGCTTACAGCTCTGTC 720  
 Db 661 TTGTCCTCTGGCGCTGGAGTCCCGGCTGTCAGGCTTACAGCTCTGTC 720

RESULTS  
 A32203 LOCUS A32203 3750 bp DNA linear PAT 10-DEC-1996  
 DEFINITION NANBH PT polyprotein fragment.  
 ACCESSION A32203  
 VERSION A32203.1 GI:1926564  
 KEYWORDS SOURCE Non-A, non-B hepatitis virus  
 ORGANISM Non-A, non-B hepatitis virus  
 REFERENCES Viruses; unclassified viruses.  
 AUTHORS 1 (bases 1 to 3750)  
 FEATURES Patent: FR 2655990-A 22 21-JUN-1991;  
 Location/Qualifiers Location/Qualifiers  
 1. >3750 /product="partial NANBH PT polyprotein"  
 /protein\_id="NC\_002095.1"  
 /db\_xref="IGI:1926565"  
 /translation="WEGVFTGLTUDAHFLSQTQAGDNPPYLYAYAVTCARAOAPP  
 PSWQWIKCILKULTRGTPPLYRIGAVONETLTHPTPKFNMADLEVTST  
 WLUVGGVLAALAACLLTGESWIVGRILSLSGPAIVPDRREVYQFDEMEBCASHRY"

QY	721 ACTAAGCCGCTAACACCCAAATCTACCCCTCTGTAAACACCCCTGGGGATGAGTAGCC	780	Db	1801 TCGAGACTCCATGCGAACCGGAACGGATGAGCTACATCGTCACCGAC	1860
Db	721 ACTAGCCGCCTCACCAACAAATCTACCCCTCTGTAACTACATTCTGCGGGATGGGCC	780	QY	1861 CCTTCCACATCACAGCAGAGACGCTTAAGCGCAGGCTGCCAGGGGCTCCCCCTCC	1920
QY	781 CCCAAACTCGCTCCCCCAGTGCTGTTCACTTGTAGCCGGCATTTGCTGGGG	840	Db	1861 CCTTCCACATCACAGCAGAGACGCTTAAGCGCAGGCTGCCAGGGGCTCCCCCTCC	1920
QY	840	840	QY	1921 TTGGCAGCTTCAGCTACCGAGTGTGCGCTTCTCGAAGGGACATACATTAC	1980
Db	841 OCTGTTGCACTAGGCTTGGAGAGGTGTCAGTCTGGACATCTTGCGGGGATGGAGCA	900	Db	1921 TTGGCAGCTTCAGCTACCGAGTGTGCGCTTCTCGGCTTCTCGAAGGGACATACATTAC	1980
QY	900	900	QY	1981 CAAATGACTTCCCAGCGTACCTATCGAGGCAACTCTGTGGGATGAGTG	2040
Db	901 GAGAGTGCAGGGCTGCTGGCTTAAAGTCATGAGCAGGGAAATGCCCTCACCGAG	960	Db	1981 CAAATGACTTCCCAGCGTACCTATCGAGGCAACTCTGTGGGATGAGTG	2040
QY	960	960	QY	2041 GGC 2043	
Db	961 GACCTGGTTACTACTCCCTGCATCTCTCCCTGCCATCTCTCGTGGCTGCGGCTCACCGAG	1020	Db	2041 GGC 2043	
QY	1021 TGGCAGCGATACTGGTGGCGACGTGGGTGAGGGGGCTGAGTGGATGAC	1080			
Db	1021 TCGCGAGGATACTGGTGGCGACGTGGGTGAGGGGGCTGAGTGGATGAC	1080			
QY	1081 CGGCTGATAGGTTGCTCCGGGGTAAACATGTTCCCACCGACTAGTGGCAGAG	1140			
Db	1081 CGGCTGATAGGTTGCTCCGGGGTAAACATGTTCCCACCGACTAGTGGCAGAG	1140			
QY	1141 AGCGAGGCCGAGCAGCTGTTACTCGATCTCTCGACCTACTATCACCAACTGTT	1200			
Db	1141 AGCGAGGCCGAGCAGCTGTTACTCGATCTCTCGACCTACTATCACCAACTGTT	1200			
QY	1201 AGAGGCTCCAGTGAACAGGAGCTCTCCACGCCCTGTCGGGTGCGCTA	1260			
Db	1201 AGAGGCTCCAGTGAACAGGAGCTCTCCACGCCCTGTCGGGTGCGCTA	1260			
QY	1261 AGGGAGTTGGACTGATGACAGTTGGCTACTTCAGCTGGCTCCAGTCC	1320			
Db	1261 AGGGAGTTGGACTGATGACAGTTGGCTACTTCAGCTGGCTCCAGTCC	1320			
QY	1321 AGCTCTGGGATTAACCGGAGRCCTTTTCATGCCAACGTTGAGGAGG	1380			
Db	1321 AGCTCTGGGATTAACCGGAGRCCTTTTCATGCCAACGTTGAGGAGG	1380			
QY	1380 GTCGGGGGAGACGGCATATGAGGATGTCAGACCCCTGCTGGAGCACAGGAA	1440			
Db	1381 GTCGGGGGAGACGGCATATGAGGATGTCAGACCCCTGCTGGAGCACAGGAA	1440			
QY	1440	1440			
Db	1440	1440			
QY	1441 CATGTCAAAACGGTCATGAGGATGTCAGACCCCTGCTGGAGCACAGGAA	1500			
Db	1441 CATGTCAAAACGGTCATGAGGATGTCAGACCCCTGCTGGAGCACAGGAA	1500			
QY	1500	1500			
Db	1500	1500			
QY	1560 GGACATTCCTCACACGATACACCCCTGGAGCTACGGCTAACAC	1560			
Db	1560 GGACATTCCTCACACGATACACCCCTGGAGCTACGGCTAACAC	1560			
QY	1561 TATTCAGGGCGTGTGGCGCTCTAGGAGTACGGAGTTACCGGGGGGG	1620			
Db	1561 TATTCAGGGCGTGTGGCGCTCTAGGAGTACGGAGTTACCGGGGGGG	1620			
QY	1620	1620			
Db	1620	1620			
QY	1620	1620			
Db	1620	1620			
QY	1680	1680			
Db	1680	1680			
QY	1680	1680			
Db	1680	1680			
QY	1680	1680			
Db	1680	1680			
QY	1740	1740			
Db	1740	1740			
QY	1740	1740			
Db	1740	1740			
QY	1740	1740			
Db	1740	1740			
QY	1800	1800			
Db	1800	1800			
QY	1800	1800			
Db	1800	1800			
QY	1860	1860			
			FEATURES		
			source	/organism="Hepatitis C virus"	
				/mol_type="genomic RNA"	
				/strain="HCV-1b"	
				/db_xref="HCV-KI-S2"	
				/note="FN-sensitive HCV from patient 2"	
			CDS	/codon_start=1	
				/product="polyprotein"	
				/protein_id="BA040506.1"	
				/db_xref="GI:1814089"	
				/translation="MSTNPKPORKRNNRRPDKPPGGGIVGVYLLPRGPRL	
				GVRTRKRSRSPRGRPKARQPEGRAMQPYPMPLYGNGWAGLISPR	
				SRPWGPTDPRRSRNLRVDTMTGCFADLWGLTIPWCAPLGCASRALAHRVLED	
				GUNVATGNGLPGCSESFILALLSLCLTIPASAYEVNNSGYHYWHDNCNSAVYEAAD	
				MIMIFPGUVFVRBANSRWRWALPIILAKNSSVFPTTIRRHVDLVUGAACSAMY	

Best Local Similarity 92.7%; Pred. No. 0;  
Matches 1894; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy	1	TGGAGGGGCTTCAAGGCCTCACCACTGGATGCCACTTCCTGTCCAAACAG	60
Db	5001	TGGAGGGGCTTCAAGGCCTCACCACTGGATGCCACTTCCTGTCCAAACAG	5060
Qy	61	CAGGAGGAGACAATCCCCTACTCTGTGGTACAGGCTACTCTGTGGCTAGGCC	120
Db	5061	CAGGAGGAGACAATCCCCTACTCTGTGGTACAGGCTCACCACTGGCTAGGCC	5120
Qy	121	CAGGCCACCTCCATGGGATCATCTGGGAAGTGTCTCATGGCTAACGCTACT	180
Db	5121	CAGGCTCACCCTCATGGGATCATCTGGGAAGTGTCTCATGGCTAACGCTACT	5180
Qy	181	CTGGGGGCCAACCCCTCTGTATAGCTGGTGGCGGCCCTAACAAAGGGTCAACCTC	240
Db	5181	CTGCTGGGCCAACCCCTCTGTATAGCTGGTGGCGGCCCTAACAAAGGGTCAACCTC	5240
Qy	241	ACACACCCATAACCAATTACATCATGGCATGCATCTAGCCGACCTGGCTGTCAG	300
Db	5241	ACACACCCATAACCAATTACATCATGGCATGCATCTAGCCGACCTGGCTGTCAG	5300
Qy	301	AGCACCTGGGCTCTGGTGGATCTCTGGCTGGGGGCTCTGGCTGGTGTCTCATGGCTAACGCTACT	360
Db	5301	AGCACCTGGGCTCTGGTGGATCTCTGGCTGGGGGCTCTGGCTGGTGTCTCATGGCTAACGACA	5360
Qy	361	GCCAGCTGGCATCTGGTGGAGGATCATCTGGTGGGGGGCTATGTGTCGAC	420
Db	5361	GCCAGCTGGCATCTGGTGGAGGATCATCTGGTGGGGGGCTATGTGTCGAC	5420
Qy	421	AGGAAGTCTACAGGAGTCGAGATGGAGTGGGGCTGGCTGCACTCCCTAC	480
Db	5421	AGGAAGTCTACAGGAGTCGAGATGGAGTGGGGCTGGCTGCACTCCCTAC	5480
Qy	481	ATCGAGGAGGAATGGAGCTGGCTGGCTGGAGTCAGTCAGCAAAGCGCTCTGGTTGTCAG	540
Db	5481	ATCGAGGAGGAATGGAGCTGGCTGGCTGGAGTCAGTCAGCAAAGCGCTCTGGTTGTCAG	5540
Qy	541	ACAGCCACGAGCAGCAGGGAGGCCCTGCTCCCGTGCTGGAGTCAGTGCGCTGGCTGGAGCCT	600
Db	5541	ACGGCCACGAGCAGCAGGGAGGCCCTGCTCCCGTGCTGGAGTCAGTGCGCTGGAGCCT	5600
Qy	601	GAGACCTCTGGGAAACACATGGTGGACTCATCGGGATACTAGTACAGGGC	660
Db	5601	GAGACCTCTGGGAAACACATGGTGGACTCATCGGGATACTAGTACAGGGC	5660
Qy	661	TTCGCCCTCGCCCTGGAGTCCCGATGTCATCATCTGATGGCTTACAGCCTCTGTC	720
Db	5661	TTCGCCCTCGCCCTGGAGTCCCGATGTCATCATCTGATGGCTTACAGCCTCTGTC	5720
Qy	721	ACTAGCCGCTTACCCCAATCTACCTCTGTGTAACATCCCTGGGGATGGTAGCC	780
Db	5721	ACTAGCCGCTTACCCCAATCTACCTCTGTGTAACATCCCTGGGGATGGTAGCC	5780
Qy	781	GCCAACTCTGGCTCCCAAGCTGGCTGCTCACTTGAGGCGCCGATGCTGGCG	840
Db	5781	GCCAACTCTGGCTCCCAAGCTGGCTGCTCACTTGAGGCGCCGATGCTGGCG	5840
Qy	841	GCTGTTGGCAGCTAGGCTTGGAGGTGGCTTGAGCTCTTGCGGCTATGGCA	900
Db	5841	GCTGTTGGCAGCTAGGCTTGGAGGTGGCTTGAGCTCTTGCGGCTATGGCA	5900
Qy	901	GGAGCTGGCAGGGGCTCGCTTGGAGGTGGCTTGAGCTCTTGCGGCTATGGCA	960
Db	5901	GGAGCTGGCAGGGGCTCGCTTGGAGGTGGCTTGAGCTCTTGCGGCTATGGCA	5960
Qy	961	GACCTGGTAACTTCTCTGCACTCTCTCTGGCTGGGGATGCTG	1020
Db	5961	GACCTGGTAACTTCTCTCTGGCTGGGGATGCTG	6020
Qy	1021	TGGCAGCGACTSGCTCGCCACGGGGCTGGCTGAGTGGCTGAGTGGTAC	1080

ORIGIN

Query Match

		LOCUS	DEFINITION	LINEAR
		Hepatitis C virus genomic RNA, complete genome, isolate:HCV1094.	ACCESSION	VRL 22-AUG-2002
		AB049088	VERSION	AB049088.1
		REFERENCE	AUTHORS	KEYWORDS
		Mishiro, S.	Hepatitis C virus	
Db	6021	TGCGCGCAAACTACTGGTCGCATGPGGCCGGAGAGGGCTGCACTGGATGAC	6080	Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
Qy	1081	CGGCTGATAGGGTCTGCCTCCGGGTAACCATGTTCCCCAACGACTATGTGCAGAG	1140	
Db	6081	CGGCTGATAGGGTCTGCCTCCGGGTAACCATGTTCCCCAACGACTATGTGCAGAG	6140	
Qy	1141	AGCGAGCCGAGCAGCAGTGACTCAGATCTCTCGACACTACTATGATGAGTG	1200	
Db	6141	AGCGAGCCGAGCAGCAGTGACTCAGATCTCTCGACACTACTATGATGAGTG	6200	
Qy	1201	AGAGAGCTCCCAAGCTGGATTACGAGGACTCTCACGCTGCGCTCGTGCTA	1260	
Db	6201	AGAGAGCTCCCAAGCTGGATTACGAGGACTCTCACGCTGCGCTCGTGCTA	6260	
Qy	1261	AGGATGTTGGACTGGATGAGGACTCTCACGCTGCGCTCGTGCTA	1320	
Db	6261	AGGATGTTGGACTGGATGAGGACTCTCACGCTGCGCTCGTGCTA	6320	
Qy	1321	AGCTCTGGCGGAATACCGAGCTCCCTTCTCATCCACGCTGCTACAGGG	1380	
Db	6321	AGGTCTCTGGCGGAATACCGAGCTCCCTTCTCATCCACGCTGCTACAGGG	6380	
Qy	1381	GCTCTGGGGAGACGGCATATGCGAGACCCCTGCTCATGGAGCACAGATACCGA	1440	
Db	6381	GCTCTGGGGAGACGGCATATGCGAGACCCCTGCTCATGGAGCACAGATACCGA	6440	
Qy	1441	CATGTCAAAMAGGTTCATGGATGTTGGCTAAGACAGCTGTGTAACATGGCAT	1500	
Db	6441	CATGTCAAAMAGGTTCATGGATGTTGGCTAAGACAGCTGTGTAACATGGCAT	6500	
Qy	1501	GGAACTTCCCATCACGCTACACACGGGCCCCCTGCAACCCCTCCGGCAAC	1560	
Db	6501	GGAACTTCCCATCACGCTACACACGGGCCCCCTGCAACCCCTCCGGCAAC	6560	
Qy	1561	TATTCCAGGGCTGGGGGTGCTGGACTGAGGACTGAGGACTGAGGAGCTGG	1620	
Db	6561	TATTCCAGGGCTGGGGGTGCTGGACTGAGGACTGAGGAGCTGG	6620	
Qy	1621	GAATTCACTACGTGAGGAGGACTGAACTGAACTAATGCCGTGAGGGTCA	1680	
Db	6621	GAATTCACTACGTGAGGAGGACTGAACTGAACTAATGCCGTGAGGGTCA	6680	
Qy	1681	GCCCCCGAATTCTCACAGAGCTGGATGGGGGGGGGGGGGGGGGGGG	1740	
Db	6681	GCCCCCGAATTCTCACAGAGCTGGATGGGGGGGGGGGGGGGGGGGG	6740	
Qy	1741	AAACCTCTCTTACGGAGGAGGAGTCACTTCCAGGGGGGGGGGGGGGG	1800	
Db	6741	AAACCTCTCTTACGGAGGAGTCACTTCCAGGGGGGGGGGGGGGGGG	6800	
Qy	1801	TGGAGCTCCCAGGAGCCGACCGATAGGGCTACTCTCATCTCACCAC	1860	
Db	6801	TGGAGCTCCCAGGAGCCGACCGATAGGGCTACTCTCATCTCACCAC	6860	
Qy	1861	CCCTCCACATCACAGAGAGGGCTAGGGCTAGGGCTAGGGCTAGGG	1920	
Db	6861	CCCTCCACATCACAGAGAGGGCTAGGGCTAGGGCTAGGGCTAGGG	6920	
Qy	1921	TGGCCAGCTCTCAGCTGGAGCTGCTCTCGCCCTCTCTGAGAGGACT	1980	
Db	6921	TGGCCAGCTCTCAGCTGGAGCTGCTCTCGCCCTCTCTGAGAGGACT	6980	
Qy	1981	CAAATGACTCCAGAGCTGACTCATCGAGCCACCTCTGCGCGCATGAG	2040	
Db	6981	CAAATGACTCCAGAGCTGACTCATCGAGCCACCTCTGCGCGCATGAG	7040	
Qy	2041	GGC 2043		
Db	7041	GGC 7043		

RESULT 9

AB049088



6693 ACCCCGGATTCCTCCGGAGGAGTGGAGTGGCTGCGAACAGTAGCTCCGGCTC 6752  
 /translation="MSTINPKQPKRKRTKNTNRDODVKEPGQGVIGVYLRPGRGL  
 GVRARAKTSESOQSRKRRNGKVIDLTFGFADJMGTPVGAPEGGAAKALANGVRVLED  
 SRPSWGPIDPRRSRNGKVIDLTFGFADJMGTPVGAPEGGAAKALANGVRVLED  
 Qy 1741 AAACCTCTCTAGGGAGGAGGTCACTTCAAGTCAGTCAGCTGCTAACAAATACCTGGTGG 1800  
 Db 6753 AAACCCCTCTACTGGAGGAGGTCACTTCAAGTCAGTCAGCTGCTAACAAATACCTGGTGG 6812  
 1801 TCCGACTCCATGGAGCCGACGGATGAGGAGCTACTTCATGTCACCCAC 1860  
 Db 6813 TCACGTCCTCATGTCAGCCGACGGATGAGGAGCTACTTCATGTCACCCAC 6872  
 1861 CCCCTCCCACATCACAGCAGAGACGGCTAAGCCAGCTGGCAGCTGGCAGGCTGCCAGGGCTCCCTCC 1920  
 Db 6933 TTGGCAGCTCTTCACTGGAGGAGGTCACTTCAAGTCAGTCAGCTGCTAACAAATACCTGGTGG 6992  
 Qy 1981 CAAATGACTCCAGAGCTGACTCATGAGCCACCTCTGGGGATGAGTG 2040  
 Db 6933 CATCTGACTCCCGAGGAGCTGACTCATGAGCCACCTCTGGGGATGAGTG 7052  
 Qy 2041 GGC 2043  
 Db 7053 GGC 7055

RESULT 10  
 AF333324 AF333324 9587 bp mRNA linear VRL 18-DEC-2002  
 DEFINITION Hepatitis C virus type 1b polyprotein mRNA, complete cds.  
 ACCESSION AF333324  
 VERSION AF333324.1 GI:12831192  
 KEYWORDS Hepatitis C virus type 1b  
 SOURCE ORGANISM Hepatitis C virus type 1b  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 Hepacivirus.  
 REFERENCE 1.(bases 1 to 9587)  
 Thomon,M., Nascimbeni,M., Gonzales,S., Murthy,K.K., Rehermann,B.  
 and Liang,T.J.  
 Emergence of a distinct pattern of viral mutations in chimpanzees  
 infected with a homogeneous inoculum of hepatitis C virus  
 Gastroenterology 121 (5), 1226-1233 (2001)  
 MEDLINE 21534507  
 PUBLMED 11677216  
 REFERENCE 2.(bases 1 to 9587)  
 Thomson,M. and Liang,J.  
 AUTHORS Title  
 JOURNAL Direct Submission  
 Submitted (03-JAN-2001) LDS, NIDDK/NIH, 10 Center Drive, Bethesda,  
 MD 20892, USA  
 FEATURES source  
 1. -9587 location/Qualifiers  
 /organism="Hepatitis C virus type 1b"  
 /mol\_type="mRNA"  
 /strain="chimera of HCV-BK, HCV-J, HCV-H"  
 /serotype="1b"  
 /specific\_host="chimpanzee"  
 /db\_xref="nt;taxon:31647"  
 /clone="HCV-CG1b"  
 5'UTR  
 1. .341  
 misc\_feature  
 1. /note="derived from BK strain; similar to sequence  
 deposited in GenBank Accession Number M58335"  
 misc\_feature  
 272..3030  
 /note="derived from J strain; similar to sequence  
 deposited in GenBank Accession Number D90208"  
 342..9374  
 /codon\_start=1  
 /product="polyprotein"  
 /protein\_id="AK08509.1"  
 /ab\_xref="GI:12831193"

Query Match 88.2%; Score 1801.4; DB 14; Length 9587;  
 Best Local Similarity 92.6%; Pred. No. 0; Mismatches 151; Indels 0; Gaps 0;  
 Matches 1892; Conservative 0; Mismatches 151; Indels 0; Gaps 0;  
 ORIGIN  
 1 TGGAGGGCTCTCACAGCTTACCCAGCTGGATGCCACCTCTGCCAACAAAG 60  
 Qy 1 5013 TGGAGAGCTCTCACAGCTTACCCAGCTGGATGCCACCTCTGCCAACAAAG 5072  
 Db 61 CAGGGAGGACAATCTCCATTCTGGTGGTACAGCTACTGTGCGCTAGGCC 120  
 Qy 5073 CAGGGAGGACAATCTCCATTCTGGTGGTACAGCTAACAGCACCGTGGCCAGGGCT 5132



ORIGIN						
	Query Match	Score	DB	Length	Indels	Gaps
Best Local Similarity	88.0 %;	Score 1798.2;	DB 14;	Length 9379;		
Matches 1890; Conservative	92.5 %;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1 TGGAGGGGCTCTCACAGGCCTCACCCAGTGATGCCACTTCTGTCCCAAACAAAG 60					
Db	5001 TGGAGGAGTCCTTAAGGCTGACCCATAGATGCCACTTCTGTCCCAAACAAAG 5060					
Qy	61 CAGGAGGAGACAATCTCCCTACTCTGGTGGGTACAGCGTACTGTGNGCCTAGGCC 120					
Db	5061 CAGGAGGAGATAACTTCCCTACTCTGGTGGGTACAGCGTACTGTGNGCCTAGGCC 5120					
Qy	121 CAGGCCAACCTCCATCATGGATCAATGTGGAGTGCTCATAGCTAANGCTACT 180					
Db	5121 CAGGCTTCACTCTCATGGGAGGACAAATGTGGAGTGCTCATAGCTAANGCTACT 5180					
Qy	181 CTGCGCGGCGCACACCCCTGCTGTATAGCTGGAGGCCGCTTACAGGCACCCCTC 240					
Db	5181 CTGCGCGGCGCACACCCCTGCTGTATAGCTGGAGGCCGCTTACAGGCACCCCTC 5240					
Qy	241 ACACACCCATAACCAATTCAATGATGATGATGATGATGATGATGATGATGATG 300					
Db	5241 ACACACCCATAACCAATTCAATGATGATGATGATGATGATGATGATGATGATG 5300					
Qy	361 GGCACCTGGGTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420					
Db	5361 GGCACCTGGGTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5360					
Qy	421 AGGGAGTCTCTACCAAGGAGTTGATGAGTGAATGAGTGAAGTGAAGTGAAGT 480					
Db	5421 AGGGAGTCTCTACCAAGGAGTTGATGAGTGAATGAGTGAAGTGAAGTGAAGT 5480					
Qy	481 ATCGAGGAGGAATGAGCTGGCGAGCTGGCGAGCTGGCGAGCTGGCGAGCTGG 540					
Db	5481 ATCGAGGAGGAATGAGCTGGCGAGCTGGCGAGCTGGCGAGCTGGCGAGCTGG 5440					
Qy	541 ACAGGCCACAGCAAGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600					
Db	5541 ACAGGCCACAGCAAGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5600					
Qy	601 GAGACCTCTGGGAACACATGATGGAATCTACAGGGGATACACTACTAGGAGG 660					
Db	5601 GAGACCTCTGGGAACACATGATGGAATCTACAGGGGATACACTACTAGGAGG 5660					
Qy	661 TTGTCACCTCTGGGAATCCCGGGATGATGATCAATGGGTACAGGCTCTGGTC 720					
Db	5661 TTGTCACCTCTGGGAATCCCGGGATGATGATCAATGGGTACAGGCTCTGGTC 5720					
Qy	721 ACTAGCCCCTCACCACTTACCCCTCTCTGGCTAACATCCTGGGGATGGTGTGAC 780					
Db	5721 ACTAGCCCCTCACCACTTACCCCTCTCTGGCTAACATCCTGGGGATGGTGTGAC 5780					
Qy	781 GCCCAACTCTCCCCCAAGTGTCTCTGGCTAACATCCTGGGGATGGTGTGAC 840					
Db	5781 GCCCAACTCTCCCCCAAGTGTCTCTGGCTAACATCCTGGGGATGGTGTGAC 5840					
Qy	841 GCTGTGGAGCATGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 900					
Db	5841 GCTGTGGAGCATGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5900					

QY	901	GAGCTGGCAGCCGCGCTCGGCCCTTAAGTCATGAGCGGAATGCCCTCACCGAG	960
Db	5901	GGGGTGGCTGGCGCACTCTGGCTTTAAGTCATGAGGGAGATGCCCTCACCGAG	5360
QY	961	GACCTGGTAATCTGCTCCCTGCATCCCTCTCTGGGCCCTGGCTCGGGTCGCG	1020
Db	5961	GACCTGATAATCTGCTCCCTGCATCCCTCTCTGGGCCCTGGCTCGGGTCGCG	6020
QY	1021	TSGCTGAGGAACTCTACTCTGCTGCACTCCCTCTGGCTCCCTGGCTCGGGTCGCG	1080
Db	6021	TSGCTGAGGAACTCTACTCTGCTGCACTCCCTCTGGCTCCCTGGCTCGGGTCGCG	6080
QY	1081	CGCTGATAGCGTGCCTCGCGGGGAACTGTCAGCTGCGGGGGCTGCTCGGGTCGCG	1140
Db	6081	CGCTGATAGCGTGCCTCGCGGGGAACTGTCAGCTGCGGGGGCTGCTCGGGTCGCG	6140
QY	1141	AGGAGCGCCGAGCGAGCGAGCTGCTGCGGGGGCTGCTCGGGTCGCG	1200
Db	6141	AGGAGCGCCGAGCGAGCGAGCTGCTGCGGGGGCTGCTCGGGTCGCG	6200
QY	1201	AAGAGGCTCACCGAGGTGATTACCGAGACTGCTCACCTGAGCTCTCCCGGGCTG	1260
Db	6201	AAGAGGCTCACCGAGGTGATTACCGAGACTGCTCACCTGAGCTCTCCCGGGCTG	6260
QY	1261	AGGGATGTTGGGACTGGATATGCACAGTTGGTACTCTAACATCACTACAGTG	1320
Db	6261	AGGGATGTTGGGACTGGATATGCACAGTTGGTACTCTAACATCACTACAGTG	6320
QY	1321	AAGCTCTCGCGGATTAACGGGAGTCCCTTCTCTAGGCCAAAGTGGTACAAGGG	1380
Db	6321	AAGATCATCCGGGTTACCGGGAGTCCCTTCTCTAGGCCAAAGTGGTACAAGGG	6380
QY	1381	GTCTGGCGGGAGACGGATCATGGAGACCGCTGCTCATGGAGACAGAACCGGA	1440
Db	6381	GTCTGGCGGGAGACGGATCATGGAGACAGAACCGGA	6440
QY	1441	CATGCTAAACCGTTCCATGAGGATCGTGGCCCTAGACCTGAGCATGGCAT	1500
Db	6441	CATGCTAAACCGTTCCATGAGGATCGTGGCCCTAGACCTGAGCATGGCAT	6500
QY	1501	GGGACATTCCCCATCAACCGATCACCAACGGGCCCTGACGGCCCTCCCAAGGCCAAC	1560
Db	6501	GGGACATTCCCCATCAACCGATCACCAACGGGCCCTGACGGCCCTCCCAAGGCCAAC	6560
QY	1561	TATTCGAGGGCGCTGGGGGGCTGTGAGGAGCTGAGGAGTACGGCGGGGG	1620
Db	6561	TATTCGAGGGCGCTGGGGGGCTGTGAGGAGTACGGCGGGGG	6620
QY	1621	GATTCCACTACGTGAGGACATGCCACTGACAACTGACAGTAATCCCCTGCAAGTCCA	1680
Db	6621	GATTCCACTACGTGAGGACATGCCACTGACAACTGACAGTAATCCCCTGCAAGTCCA	6680
QY	1681	GCCCGGAACTCTGAGAGTGGATCGGGTCCGGCTGCGCTGAGGAGCTCCCGGTC	1740
Db	6681	GCCCGGAACTCTGAGAGTGGATCGGGTCCGGCTGCGCTGAGGAGCTCCCGGTC	6740
QY	1741	AAACCTCTCTAGGGGAGGGGCAATTCGAGGGTGGTACAGGAGCTCCCGGTC	1800
Db	6741	AAACCTCTCTAGGGGAGGGGCAATTCGAGGGTGGTACAGGAGCTCCCGGTC	6800
QY	1801	TCCGAGCTCCATCGAGCCGACGGATGAGCTCACTCCATGCTCACCGAC	1860
Db	6801	TCCGAGCTCCATCGAGCCGACGGATGAGCTCACTCCATGCTCACCGAC	6860
QY	1861	CCCTCCACATCACAGAGAGACGGCTAAGCGAGCTGGCTGGGGCTCCCTCC	1920
Db	6861	CCATCCACATCACAGAGAGACGGCTAAGCGAGCTGGCTGGGGCTCCCTCC	6920
QY	1921	TTCGCACTCTGAGCTAGGAGCTGCTGCGCTCTGGAGGAGCATATTAC	1980
Db	6921	TTCGCACTCTGAGCTAGGAGCTGCTGCGCTCTGGAGGAGCATATTAC	6980
QY	1981	CAAATGACTTCCAGACCTGACTCATCGAGGCCAACCTCCNGTGGGGAGAGATG	2040
Db	6981	CATCATGACTCCCGAGCTGACCTCATCGAGGCCAACCTCTGGGGAGAGATG	7040
QY	2041	GGC 2043	
Db	7041	GGC 7043	
RESULT	12		
AR005087	AR005087		
Locus	Sequence 35 from patent US 5747339.		
DEFINITION	Sequence 35 from patent US 5747339.		
VERSION	AR005087.1		
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 7863)		
AUTHORS	Okayama, H., Fukui, I., Mori, C., Takamizawa, A. and Yoshida, I.		
TITLE	Non-A, non-B hepatitis virus genomic CDNA and antigen polypeptide		
JOURNAL	Patent: US 5747339-A 35 05-MAY-1998;		
FEATURES	Location/Qualifiers		
source	1. - 7863 /organism="unknown" /mol_type="unassigned DNA"		
ORIGIN			
Query	Match		
Best local similarity 92.5%; Pred. No. 0;			
Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;			
Db	3505 TGGGAGAGCTCTCACGGGCTCACCCCTATATGCGACATCTCTGCGCACACGAC	3564	
QY	61 CAGGAGGAGAACTCCCTACGGGGTACAGGCTCACCCAGCTGCGCTAGGCCAACCTGGCTAGGGCC	120	
Db	3565 CAGGAGGAGAACTCCCTACGGGGTACAGGCTCACCCAGCTGCGCTAGGCCAACCTGGCTAGGGCC	3624	
QY	121 CAGGCCAACCTCCATCATCGTGGATCAATGTGGAGTCTCATAGGGTAACCTACT	180	
Db	3625 CAGGCCAACCTCCATCATCGTGGATCAATGTGGAGTCTCATAGGGTAACCTACT	3684	
QY	181 CTGGGGGGGAAACCCCTGCTGAGGAGCTGAGGAGTACGGCGGGTCCAAACAGGGTACCGGTCACCC	240	
Db	3685 CTGCAAGGACCAACCCCTGCTGAGGAGTACGGCGGGTCCAAACAGGGTACCGGTCACCC	3744	
QY	241 ACACACCCATACCAATTCATCATGGATGATGTCAGCGACCTGGAGGTCCTACG	300	
Db	3745 ACCACACCCATACCAATTCATCATGGATGATGTCAGCGACCTGGAGGTCCTAC	3804	
QY	301 AGCACCTGGGTGCTGGTGGGGGGGCTCTGCACTCTGGCTGAGTATGGCTGACACA	360	
Db	3805 AGCACCTGGGTGCTGGTGGGGGGCTCTGCACTCTGGCTGAGTATGGCTGACACA	3864	
QY	361 GGCGAGGAGCTGATGTTGGTGGAGTCACTTGTGGGGGGGTTGGTCCGAC	420	
Db	3865 GGCGAGGAGCTGATGTTGGTGGAGTCACTTGTGGGGGGGTTGGTCCGAC	3924	
QY	421 AGGGAGCTCTACGAGGAGTGGATGAGGAGAGTGGCGGTGCGACCTCC	480	
Db	3925 AGGGAGCTCTACGAGGAGTGGATGAGGAGAGTGGCGGTGCGACCTCC	3984	
QY	481 ATCGAGGAGGAACTGAGGCTGCGAGGAGTCAAGGAGAAGGCTCGGGTCTGCG	540	
Db	3985 ATCGAGGAGGAACTGAGGCTGCGAGGAGTCAAGGAGAAGGCTCGGGTCTGCG	4044	
QY	541 ATCGAGGAGGAACTGAGGCTGCGAGGAGTCAAGGAGAAGGCTCGGGTCTGCG	600	
Db	4045 ATCGAGGAGGAACTGAGGCTGCGAGGAGTCAAGGAGAAGGCTCGGGTCTGCG	4104	

Qy	601	GIGACACTCTGCGCGAACACATGTGGACTTCATACGGGATACTGACTTAGCGGC	660	Db	
Db	4105	GAGACATTCCTGCGCGAACACATGTGGACTTCATACGGGATACTGACTTAGCGGC	4164	Qy	
Qy	661	TGTC CACTCTGCGGGAAATCCGGATTGCACTCATGAGGGATACTGACTTAGCGGC	720	Db	
Db	4165	TTATC ACTCTGCGGGAAACCCGCAATAGCATATGATGCCATCACGCTCTTCATC	4224	Qy	
Qy	721	ACTAGCCGCTTACACCCAATCTAACCTCTGCTTAACATCCTGGGGTACAGCTCTGTC	780	Db	
Db	4225	ACCA GCGCTCACCAACCAAGTAACCTCCCTGTTAACATCTGCGGGTGGCT	4284	Qy	
Qy	781	GCCCACTCGCTCCCCAGTGCTCAGTTCTGAGGCCGATCTGCGGCC	840	Db	
Db	4285	GCCTCACTCGCCCCCAGCGCGCCTCGGGCTTGGAAGGGTCTGCTGAGGCTCG	4344	Qy	
Qy	841	GCTGTGCGAGCATAGGGCTTGGAAGGGTCTGCTGAGGCTCG	900	Db	
Db	4345	GCTGTGCGAGCATAGGGCTTGGAAGGGTCTGCTGAGGCTCG	4404	Qy	
Qy	901	GGAGTGGCGCGCGCTGCTGCGCTTAAAGTCATGAGGCTCGGCCAGATGCCCTCA	960	Db	
Db	4405	GGAGTGGCGCGCGCTGCTGCGCTTAAAGTCATGAGGCTCGGCCAGATGCCCTCA	4464	Qy	
Qy	961	GACCTGGTTACTACTCCTGCATCTCTCTCTCTGATGCCCTGCGTCGGGTCTG	1020	Db	
Db	4465	GACCTGGTCAATCTACTCCTGCATCTCTCTCTGCGCTGCGTCGGGTCTG	4524	Qy	
Qy	1021	TGCGCAGGATACCTGCGTGGCACTGGTGTGCAATGCCAGGGGCTGTCAGTGGAC	1080	Db	
Db	4525	TGCGCAGGATACCTGCGTGGCACTGGTGTGCAATGCCAGGGGCTGTCAGTGGAC	4584	Qy	
Qy	1081	C CGCTGATAGCGTCCGCTCGGAGGTAACTGGTGTGCACTGGTGTGCAAG	1140	Db	
Db	4585	CGGCTGATAGCGTCCGCTCGGAGGTAACTGGTGTGCACTGGTGTGCAAG	4644	Qy	
Qy	1141	AGGAGCGCGAGCGAGGTTCACTCGATCTCCGACTTACTATACCGACTGTG	1200	Db	
Db	4645	AGCGAGCGCGAGCGAGGTTCACTCGATCTCCGACTTACTATACCGACTGTG	4704	Qy	
Qy	1201	AAGAGGCTTACCAAGGTGTTAACAGGGAGGTGTCAACGCCCTGCTCGGTGTTGGCTA	1260	Db	
Db	4705	AAGAGGCTTACCAAGGTGTTAACAGGGAGGTGTCAACCGCTGCTCGGTGTTGGCTA	4764	Qy	
Qy	1251	ASGGATGTTGGACTGGAATGCAAGCTTGGTACTCAAGACCTGGCTCCAGRC	1320	Db	
Db	4765	AGGGATGTTGGACTGGAATGCAAGCTTGGTACTCAAGACCTGGCTCCAGRC	4824	Qy	
Qy	1321	AAGCTCTCGCGGATTAACGGGAGGCCCTTTCTGATGCCAACGGGGTACAGGG	1380	Db	
Db	4825	AAGCTCTCGCGGAGCTACTGGTGTGCACTGGCTTCTGTCATGCCAACGGGTACAGGG	4884	Qy	
Qy	1381	GTCAGCGCGAGGAGGGCATCGAGAACCTGCTCATGAGCACAGATCACCGA	1440	Db	
Db	4885	GTCAGCGCGAGGAGGGCATCGAGAACCTGCTCATGAGCACAGATCACCGA	4944	Qy	
Qy	1441	CATGCTAAACAGGTCCATGAGGATCGTGGCTACATGTCATGCGCAT	1500	Db	
Db	4945	CATGCTAAACAGGTCCATGAGGATCGTGGCTACATGTCATGCGCAT	5004	Qy	
Qy	1501	GGAACTTCCCATAACCGATACACCGACGGCCCTGAGCCCTCCAGGCCAAC	1560	Db	
Db	5005	GGAACTTCCCATAACCGATACACCGACGGCCCTGAGCCCTCCAGGCCAAC	5064	Qy	
Qy	1561	TATTCGAGGCGCTGTGGGGTGTGAGGATAGCTGGAGGTAGCGGGTGGGG	1620	Db	
Db	5065	TATTCGAGGCGCTGTGGGGTGTGAGGATAGCTGGAGGTAGCGGGTGGGG	5124	Qy	
Qy	1621	GATTCCACTACGAGGAGCATGCCACTGACACGTAATCCCGTCCAGGTCCA	1680	Db	
Db	5125	GATTCCACTACGAGGAGCATGCCACTGACACGTAATCCCGTCCAGGTCCA	5184	Qy	
Qy	1681	GCCCCGGAATCTTCACAGAAGTGGATGGGGTGGCTGCGCTCACAGGTAGCTCGGGTGG	1740	Db	
Qy	5185	GCTCTGATTTCTCGAGGCTGAGCGAGCTGGTGGCAAGGTAACCTCGCTC	5244	Db	
Qy	1741	AAACCTCTCTCGGGAGGTCACATCCTCCAGCTGAGCTACACATCTGTC	1800	Db	
Db	5245	AGGCTCTCTACGGAGGAGTACATCTCCAGCTGAGCTACACATCTGTC	5304	Qy	
Qy	1801	TGGAGCTCCAAGGAGCCGACCGATGAGCTACTTCATGATGCTACCGAC	1860	Db	
Db	5305	TACAGCTACCTGCGAGCCACCGCTACTTCATGCTACCGAC	5364	Qy	
Qy	1861	CCPCCACATCACAGAGAGGCTGCGGCCATCCGGTGG	1920	Db	
Db	5365	CCCTCCACATCACAGAGAGGCTGCGGCCATCCGGTGG	5424	Qy	
Qy	1921	TGCGCAGCTCTGAGCTGGCTGCTCTCTCGAGGACATACATTC	1980	Db	
Db	5425	TGCGCAGCTCTGAGCTGGCTGCTCTCTCGAGGACATACATTC	5484	Qy	
Qy	1981	CAAATGACTTCAGAGCTGACTCTCTGAGGCCACCTCTGGGGTGTGAGG	2040	Db	
Db	5485	CACCATGCTCGGGAGCTACCTCATCGAGGCCACCTCTGAGGGAGGATG	5544	Qy	
Qy	2041	GCG 2043	2043	Db	
Db	5545	GCG 5547	5547	Qy	
Qy	RESULT 13			Db	
Db	AR064526	AR064526	AR064526	Qy	
LOCUS	Sequence 35 from patent	7863 bp	DNA	Db	
DEFINITION	ACCESSION	35	LINEAR	PAT 29-SEP-1999	
VERSION	AR064526	AR064526	7863	bp	
KEYWORDS	AR064526.1	GI:5993834	7863	bp	
SOURCE	Unknown.	Unknown.	7863	bp	
ORGANISM	Unclassified.	Unclassified.	7863	bp	
REFERENCE	1 (bases 1 to 7863)	Okayama,H., Fukui,I., Mori,C., Takamizawa,A. and Yoshida,I.	7863	bp	
AUTHORS			7863	bp	
TITLE		Non-A, non-B hepatitis virus genomic cDNA and antigen polypeptide	7863	bp	
JOURNAL		Patent: US 584701-A 35 08-DEC-1998;	7863	bp	
FEATURES	Location/Qualifiers	Location/Qualifiers	7863	bp	
source	1. .:863	/organism="unknown"	7863	bp	
ORIGIN	/mol_type="unassigned DNA"		7863	bp	
Query	Match	87.9%	Score	1796.6	
Best	Local Similarity	92.5%	Pred. No.	0;	
Matches	1889; Conservatve	0;	Mismatches	154;	
		Indels	0;	Gaps	0;
Db	4825	AAGCTCTCGCGAGCTACTGGTGTGCACTGGCTTCTGTCATGCCAACGGGTACAGGG	4884	Qy	
Qy	1	TGGAGGGCTCTGACAGGCCCTACCCACGCTGAGTGTGCCACCTCTGTCACAGGAA	60	Db	
Db	3505	TGGAGAGTGTCTCACAGGCCCTACCCATAGTGCACCTCTGTCACAGGAA	3564	Qy	
Qy	61	CAGCCAGGAGACACTCCCTACCTGGCTTACAGGCTCTGTCCTAGGG	120	Db	
Db	3565	CAGGAGGAGACACTCCCTACCTGGCTTACAGGCTCTGTCCTAGGG	3624	Qy	
Qy	121	CAGGGCCACCTCCATGGATCAATGGAGCTCTATACGCTTAAGCTT	180	Db	
Db	3625	CAGGGCCACCTCCATGGATCAATGGAGCTCTATACGCTTAAGCTT	3684	Qy	
Qy	181	CAGGGCCACCTCCATGGATCAATGGAGCTCTATACGCTTAAGCTT	240	Db	
Db	3685	CAGGGCCACCTCCATGGATCAATGGAGCTCTATACGCTTAAGCTT	3344	Qy	
Qy	241	ACACACCCATACCAATCTCATGCGATCATGTCAGGCCACCTGGAGTGTGCA	300	Db	
Db	3745	ACCCACCCATACCAATCTCATGCGATCATGTCAGGCCACCTGGAGTGTGCA	3804	Qy	
Qy	301	ACGACCTGGGTGTGGGGGGCTCTGCAAGGCTCCAGGTTAGTCACCC	360	Db	

Db	4885	GCTGGGGGGAGACGCCATGCAACACCTGCCATGTGGACAGATCACCGGA	4944
Qy	1441	CATGTCAAAACGGTTCATGAGGATGTTGGCCATAGACTGTGTCAT	1500
Db	4945	CATGTCAAAACGGTTCATGAGGATGTCGGCCATAGACTGTGTCAT	5004
Qy	1501	GGACATCCCCATCAAAGTACACACAGGGCCCTTGCAAGCCCTCCAGGCCAAC	1560
Db	5005	GGACATCCCCATCAAAGTACACACAGGGCCCTTGCAACAGGCCAAC	5064
Qy	1561	TATCCAGGGCTGTGGGGGTGCTGTGAGGATACGGGTTACGGGGTGGG	1620
Db	5065	TATCCAGGGCTGTGGGGGTGCTGTGAGGATACGGGTTACGGGGTGGG	5124
Qy	1621	GATTCACTACTGTGAGGAGTACCAACGTAATGCCCSCAGTTCCA	1680
Db	5125	GATTCACTACTGTGAGGAGTACCAACGTAATGCCCSCAGTTCCA	5184
Qy	1681	GCCCCGAAATTTCAGAGAAGTGTAGGGGGCAGCTGCACAGTAGCAGG	1740
Db	5185	GCCCCGAAATTTCAGAGAAGTGTAGGGGGCAGCTGCACAGTAGCAGG	5244
Qy	1741	AACACTCTTAAGGGAGGTACATTCAGGTTGGGGTCAACAAATCTGGTGG	1800
Db	5245	AACACTCTTAAGGGAGGTACATTCAGGTTGGGGTCAACAAATCTGGTGG	5304
Qy	1801	TGCGAGTCCCATGCGGCCAACGGATGAGCTGTCACTTCATGTCACCGAC	1860
Db	5305	TGCGAGTCCCATGCGGCCAACGGATGAGCTGTCACTTCATGTCACCGAC	5364
Qy	1861	CCCTCCACATCACAGAGACGGCTTAAGGAGTTACCTCCAGGTTGGGCTCAACAAATCTGGTGG	1920
Db	5365	CCCTCCACATCACAGAGACGGCTTAAGGAGTTACCTCCAGGTTGGGCTCAACAAATCTGGTGG	5424
Qy	1921	TGGCCAGCTTCAGGTAGCCAGGTTCTGGCCCTTCCTGAAGGGACATACATTAC	1980
Db	5425	TGGCCAGCTTCAGGTAGCCAGGTTCTGGCCCTTCCTGAAGGGACATACATTAC	5484
Qy	1981	CAAATTCATTCAGGGCTGACCTCATCAGGGCAACCTTCCTGGGCCATGAGATG	2040
Db	5485	CAAATTCATTCAGGGCTGACCTCATCAGGGCAACCTTCCTGGGCCATGAGATG	5544
Qy	2041	GCGTGGGGTTAAGGTTACCTCCCTCCATCTCTCTGGGCCCTGGCTGGGTGTG	4464
Db	5545	GCGTGGGGTTAAGGTTACCTCCCTCCATCTCTCTGGGCCCTGGCTGGGTGTG	4524
<b>RESULT 14</b>			
Qy	1021	TGCGAGCATGAGTACTGCTCGACGCGTGGTCAGGGAGGGGGGTGCACTGAC	1080
Db	4525	TGCGAGCATGAGTACTGCTCGACGCGTGGTCAGGGAGGGGGGTGCACTGAC	4584
Qy	1081	CGGCTGATAGGTTCTCGGGTACCATGCTCCACGGCACTAGTGCCAGAG	1140
Db	4465	CGGCTGATAGGTTCTCGGGTACCATGCTCCACGGCACTAGTGCCAGAG	4524
Qy	1141	ACGAGCCGGAGCACGTGACTCAGATCTCCACCTACTATACCAACTGTG	1200
Db	4645	ACGAGCCGGAGCACGTGACTCAGATCTCCACCTACTATACCAACTGTG	4704
Qy	1201	AGAGGCTCCACCAGTGGATTACAGGACTCTCCACGGCTGTCGGGTGTTGCTA	1260
Db	4705	AGAGGCTCCACCAGTGGATTACAGGACTCTCCACGGCTGTCGGGTGTTGCTA	4764
Qy	1261	AGGAGGTTGGACTGATGACAGTTGGTACTCTAACCTGGCTCCAGST	1320
Db	4765	AGGAGGTTGGACTGATGACAGTTGGTACTCTAACCTGGCTCCAGST	4824
Qy	1321	AGCTCTGGCGGATGACAGTTGGTACTCTAACCTGGCTCCAGST	1380
Db	4825	AGCTCTGGCGGATGACAGTTGGTACTCTAACCTGGCTCCAGST	4884
Qy	1381	GTCTGGGGAGACGCACTATGCAACACCCTGTCTCATGGACACAGTACACCGA	1440
<b>REFERENCE</b>			
Qy	1	(bases 1 to 7863)	
Qy	1	Okanya, H., Fukui, I., Mori, C., Takamizawa, A. and Yoshida, I.	
Qy	1	Non-A, non-B hepatitis virus genomic cDNA and antigen polypeptide	
Qy	1	Patent: US 5998130-A 35 07-DEC-1999;	
Qy	1	Location/Qualifiers	
Qy	1	1..7863	
Qy	1	/organism="unknown"	
Qy	1	/mol_type="unassigned DNA"	
<b>Query Match</b>			
Qy	1	Best Local Similarity 87.9%; Score 1796; DB 6; Length 7863;	
Qy	1	Matches 1889; Conservative 0; Mismatches 154; Deletions 0; Gaps 0;	
Qy	1	TGGGAGGGCGCTTCACAGGGCTCACCCACCTGGCCACTTCCTGTCCAAACAAAG	60





Db	5425	TTGGCCAGCTTTCAGCTAGCAGTGTCTGGCCTTCCTGAAGGGACATGCACTACC	5484
Qy	1981	CAAATGACTTCCAGACGGGACCTCATCGAGGCCAACCTCTGGGCACATAGATG	2040
Db	5485	CACCATGTCCTCCGGAGCGCTCATCGAGGCCAACCTCTGGGCACAGAGATG	5544
Qy	2041	GCG 2043	
Db	5545	GCG 5547	

Search completed: February 25, 2005, 05:54:54  
Job time : 8561 sec

Run on: February 25, 2005, 01:22:04 ; Search time 957 Seconds  
 (without alignments)  
 12637.442 Million cell updates/sec

OM nucleic - nucleic search, using sw model

Title: US-09-664-363-20

Perfect score: 2043

Sequence: TGGAGGGCTCTCACAGG.....TGTGGGATAGATGGC 2043

Scoring table: IDENTITY\_NUC Gapct 10.0 , Gapext 1.0

Searched: 439206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database :	Score	Query	Length	DB	ID	Description
	1:	GeneSeq_169ec04:	*			
	2:	geneseqm1990b:	*			
	3:	geneseqm2000b:	*			
	4:	geneseqn001as:	*			
	5:	geneseqn2001bs:	*			
	6:	geneseqn2002as:	*			
	7:	geneseqn2002bs:	*			
	8:	geneseqn2003as:	*			
	9:	geneseqn2003bs:	*			
	10:	geneseqn2003cs:	*			
	11:	geneseqn2003ds:	*			
	12:	geneseqn2004as:	*			
	13:	geneseqn2004bs:	*			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	2041.4	99	9	3750	2	AHQ12241
2	2033.4	99	5	7065	2	AHQ46195
3	1803	88	3	5955	8	ACM61114
4	1803	88	3	5965	8	ACM61111
5	1803	88	3	37090	8	ACM61113
6	1801.4	88	2	7987	13	ADR38452
7	1796.6	87	2	9589	13	ADR82189
8	1796.6	87	9	9416	2	AHQ20268
9	1796.6	87	9	9416	2	AHQ21829
10	1796.6	87	9	9416	2	AAX59394
11	1787	87	5	7987	6	ADZ52321
12	1787	87	5	7989	3	AAX98968
13	1787	87	5	7989	6	ADZ52322
14	1787	87	5	7989	12	ADJ57845
15	1787	87	5	7992	6	AAL47276
16	1787	87	5	8001	2	Aaa98957
17	1787	87	5	8637	3	AAX98956
18	1787	87	5	8638	6	ABK88595
19	1787	87	5	8639	6	ABK88573
20	1787	87	5	8649	3	AAX98959

ALIGNMENTS

RESULT	ID	Description
1	AAQ12241	AAQ12241 standard; cDNA; 3750 BP.
2	AAQ12241;	
3	XX	
4	XX	
5	XX	
6	XX	
7	XX	
8	XX	
9	XX	
10	XX	
11	XX	
12	XX	
13	XX	
14	XX	
15	XX	
16	XX	
17	XX	
18	XX	
19	XX	
20	XX	

RESULT 1

ID	Description
AAQ12241	AAQ12241 standard; cDNA; 3750 BP.
PR	PR 18-DEC-1989; 89GB-00028562.
PR	PR 03-MAR-1990; 90GB-00004914.
PA	PA (WELL ) WELLCOME FOUND LTD.
PA	PA (HIGH) HIGHFIELD P. B.
XX	
PI	PI Highfield PE, Rodgers BC, Tedder RS, Barbara JAU;
XX	
DR	DR WPI; 1991-187584/26.
DR	DR P-PSD8; AARI2599.
XX	
PT	PT Post-transfusional non-A non-B hepatitis poly-peptide(s) - and also DNA and antibodies used in diagnostic assays and in vaccines.
XX	
PS	PS Claim 10; Page 88-97, 108pp; English.
XX	
CC	CC This sequence probably encodes viral non-structural proteins of the PR-NANBH viral genome which are antigenic. It was isolated from serum of humans infected with the virus. See also AAQ12236-40 and AAQ12242. (updated on 25-MAR-2003 to correct PA field.)





FT	/product= "Met-NS3-NS4A-NS4B-NS5A-NS5B"
	/note= "No stop codon shown"
QY	1321 AAGCTTCTGCGCGATTAACGGGAGTCGCTCTTTCTATGCCAACGGGTCAGGGTCAACGGG 1380
DB	3544 AAGCTTCTGCGCGATTAACGGGAGTCGCTCTTTCTATGCCAACGGTCAACGGG 3603
QY	1381 GCTGGCGGGAGACGGCATCATCAGAACCCACTCTCATGGAGCATGGAGCAAGATCACCGGA 1440
DB	3604 GCTGGCGGGAGACGGCATCATCAGAACCCACTCTCATGGAGCATGGAGCAAGATCACCGGA 3663
QY	1441 CATGTCAAACACGGTCCATGAGGATCGTGAGGCTAAGACCTGTAGAACATGGGAT 1500
DB	3664 CATGTCAAACACGGTCCATGAGGATCGTGAGGCTAAGACCTGTAGAACATGGGAT 3723
QY	1501 GGAACTTCCCATAACGCATACACCGGGCTGACCGCCCTCCAGGCCAAC 1560
DB	3724 GGAACTTCCCATAACGCATACACCGGGCTGACCGCCCTCCAGGCCAAC 3783
QY	1561 TATTCAGGGCTGTGGGGGCTGCTGAGGAGTACGGGGTACGGGGGG 1620
DB	3784 TATTCAGGGCTGTGGGGGCTGCTGAGGAGTACGGGGTACGGGGGG 3843
QY	1621 GATTTCACACTACGGGAGCATGACCACTGACAACTGAAATGCCCGTGCCAGGTCCA 1680
DB	3844 GATTTCACACTACGGGAGCATGACCACTGACAACTGAAATGCCCGTGCCAGGTCCA 3903
QY	1681 GCCCCGAACTCTACAGAGACTGATGGGGTGGCTGACAGGTACGCTCCGGCG 1740
DB	3904 GCCCCGAACTCTACAGAGACTGATGGGGTGGCTGACAGGTACGCTCCGGCG 3963
QY	1741 AAACCTCTCTACGGAGGGAGTCATCCAGGTGGCTAACAACTCTGGTGG 1800
DB	3964 AAACCTCTCTACGGAGGGAGTCATCCAGGTGGCTAACAACTCTGGTGG 4023
QY	1801 TCCGAGCTCCATGGAGGCCGACACGGGTAGGAGTGTCTCACTTCCATGTCACCCAC 1860
DB	4024 TCGAGCTCCATGGAGGCCGACACGGGTAGGAGTGTCTCACTTCCATGTCACCCAC 4083
QY	1861 CCCCTCCACATCAGCAGAGACGGCTAAGGGAGGCTCCCTCCCTCC 1920
DB	4084 CCCCTCCACATCAGCAGAGACGGCTAAGGGAGGCTCCCTCCCTCC 4143
QY	1921 TTGGCAGCTCTCAGCTAGCCAGTGTCTGGCTTCTGAGGGACATCATACC 1980
DB	4144 TTGGCAGCTCTCAGCTAGCCAGTGTCTGGCTTCTGAGGGACATCATACC 4203
QY	1981 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTCTGGGGCATGAGG 2040
DB	4204 CAAATGACTTCCAGACGCTGACCTCATGAGGCCAACCTCTGGGGCATGAGG 4263
QY	2041 GGC 2043
DB	4264 GGC 4266
RESULT 3	
ID	ACR6114 standard; cDNA; 5955 BP.
XX	AC ACR6114;
XX	AC ACR6114;
DT	09-JUL-2003 (first entry)
DE	HCV cDNA encoding Met-NS3-NS4A-NS4B-NS5A-NS5B (active).
XX	HCV; ss: gene; non-structural protein; NS3: NS4A; NS4B; NS5A; NS5B; adenoviral vector; HCV infection; vaccine; gene therapy; protease.
OS	Hepatitis C virus.
XX	Key location/Qualifiers
FT	1 .5955 a
QY	Query Match 98.3%; Score 1803; DB 8; Length 5955; Best local Similarity 92.7%; Pred. No. 0; Matches 1893; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY	1 TGGAGGAGCTTCACAGGCTCACCCAGCTGATGCCACCTCTGCGCCAAACAG 60 DB 1597 TGGAGGAGCTTCACAGGCTCACCCAGCTGATGCCACCTCTGCGCCAAACAG 1656
QY	61 CAGCAGGAGACACTCCCTACTGGGGTACCGAGCTGTGTGCGTAGGGCC 120 DB 1657 CAGCAGGAGACACTCCCTACTGGGTACCGATACCAAGCAGCTGCGCCAGGCT 1716

Qy	121	CAGGCCCACTCCATCATGGGATCAANTGGAAAGTCTCATACGGCTAACGGCTACT	180	Db	2197	AAGGGCTTACCAAGTAGTGTATTAGAAGACTGCTCCACCGTGTCCGGCTGGACTA	2856
Db	1717	CAGGCCCACTCCATCATGGGATCAANTGGAAAGTCTCATACGGCTAACGGCTACT	1776	Qy	1261	AGGATTTGGACTGATATGCAGCTTGGCTACTCAAGGCTGCTGCTCAC	1320
Qy	181	CTGGCGGGCCACACCCCTTGTTGTATGGCTGGACCGTCACAAACGAGTCACCTC	240	Db	2157	AGGATTTGGACTGATATGCAGCTTGGCTACTCAAGGCTGCTGCTCAC	2916
Db	1777	CTGGCGGGCCACACCCCTTGTTGTATGGCTGGACCGTCACAAACGAGTCACCTC	1836	Qy	1281	AACTCTTGCCGGATTACGGAGTCCCCCTTCATGCCAACGGTACAGGG	1380
Qy	241	ACACACCCATACACCAATTATCATGCGATCATGCGACTCGAGGAGTCACG	300	Db	2117	AAGCTCTGCCGAGTACCGGAGTCCCTTTCTCGTCAACGGGTAACAGGA	2976
Db	1837	ACACACCCATACACCAATTATCATGCGATCATGCGACTCGAGGAGTCACG	1896	Qy	1281	GCAGCTGGGCTCTGGGGGCTCTGGAGCTCTGGAGCTGGGACTTCAGTACACA	1440
Qy	301	AGCACCTGGGCTCTGGGGGCTCTGGAGCTCTGGAGCTGGGACTTCAGTACACA	360	Db	2977	GTCGGGGAGACGGCATATGCAACACACTGCGCATGGAGACAGCACCGGA	3036
Db	1897	AGCACCTGGGCTCTGGGGGCTCTGGAGCTGGGAGTCAGTCTGGCGCTATG	1956	Qy	1441	CATGTCACAAACGGTTCATGGAGCTGGGAGTGGCTTAACTGTTAGTGCAT	1500
Qy	361	GCAGCTGGTCAATGGGGAGATCATCTTGCTGGGGCTATTTCCGGAC	420	Db	3137	CATGTCACAAACGGTTCATGGAGCTGGGAGTGGCTTAACTGTTAGTGCAT	3096
Db	1957	GCAGCTGGTCAATGGGGAGATCATCTTGCTGGGGCTATTTCCGGAC	2016	Qy	1501	GGACATCCCACAAAGGATACACGGGGCCCTGCAACCCCTCCAGGGCAAC	1560
Qy	421	AGGGAAGTCCCTTACCAAGGATCGATGAGTGGAGTCAGCTCCCTAC	480	Db	3197	GGACATCCCACAAAGGATACACGGGGCCCTGCAACCCCTCCAGGGCAAC	3156
Db	2017	AGGAGTTCTCTACCAAGGAGTCAGTGAATGGAAAGTGGCTGGCTACTCAA	2076	Qy	1561	TATTCGGGGCTGTCGGGGCTGGCTGGCTGGAGTGGAGTACGGTGGAGT	1620
Qy	481	ATCAGGAGGGATGGAGCTCCGGAGCTCAAGGAAAGGGCTGGGTCAG	540	Db	3157	TATTCGGGGCTGTCGGGGCTGGCTGGAGTGGAGTACGGTGGAGT	3216
Db	2077	ATCAGGAGGGATGGAGCTGGCTGGAGTGGCTGGGCTGGCTGGAGTAC	2136	Qy	1621	GATTCCACTAGTGGAGCATGACCACTGACAACATAATGCCCTGCAGGTTCA	1680
Qy	541	ACAGCCACCAAGGAAAGGGGAGCCGGCTGCTCCCGTGGAGTCCAA	600	Db	3117	GATTCCACTAGTGGAGCATGACCACTGACAACATAATGCCCTGCAGGTTCA	3276
Db	2137	ACAGCCACCAAGGAAAGGGGAGCTGGCTCCGGTGGAGTCCAACTGAGC	2196	Qy	1681	GGCCCAATTCTCAAGAGTGGATGGGGGGGGGGGGGGGGGGGGGGGG	1740
Qy	601	GAGACCTCTGGGAAACACATGAGTGGAACTCATGCGGAAACAGTACAG	660	Db	3277	GGCTCTGAACTTTCAGGGAGTGGGGGGGGGGGGGGGGGGGGGGGG	3336
Db	2197	GAGACATTCCTGGGAAACACATGAGTGGAACTCATGCGGAAACAGTACAG	2256	Qy	1741	AACTCTTCTACGGGGAGGTCACATTCAGGGGGGGGGGGGGGGGGGG	1800
Qy	661	TGTCGACTCTGGGAAACACATGAGTGGAACTCATGCGGAAACAGTACAG	720	Db	3337	GGCCCTCTCTACGGGGAGGTCACATTCAGGGGGGGGGGGGGGGGG	3396
Db	2257	TPATCCACTCTCTGGGAAACACATGAGTGGAACTCATGCGGAAACAGTAC	2316	Qy	1801	TGGCAGTCCCTGCGGCCAACCGATGTCAGCTGTCACCTTCATGCGGAC	1860
Qy	721	ACTAGCCGCTAACACCAACAACTACCTCTCTGCTAACATCTGGGGGATGG	780	Db	3397	TCACTACATGCGGCCAACCGATGTCAGCTGTCACCTTCATGCGGAC	3456
Db	2317	ACAGCCGCGAACACCAACGGTACCTCTCTGTTAACATCTGGGGGGGG	2376	Qy	1861	CCCTCCACATCACAGGAGAGGGCTAAGCCAGGGCTCCCTCCCTCC	1920
Qy	781	GGCCAACTCTGGGAAACACATGAGTGGAACTCATGCGGAAACAGTACAG	840	Db	3457	CCCTCCACATCACAGGAGAGGGCTAAGCCAGGGCTCCCTCCCTCC	3516
Db	2377	GCCCAACTCGCCCCCGAGGCCGCTTCGGCTTGGGGGGGGGGGGGG	2436	Qy	1921	TGGCCGCTCTCACTAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGG	1980
Qy	841	GCTGTGGAGCATAGGCTCTGGAGGTGCTGGAGCATCTGGGGGGGGGG	900	Db	3517	TGGCCGCTCTCACTAGGCTGGGGGGGGGGGGGGGGGGGGGGGGGG	3576
Db	2437	GCTGTGGAGCATAGGCTCTGGAGGTGCTGGAGCATCTGGGGGGGGGG	2496	Qy	1981	CAAATGACTTCAGGGCTGACTCATGAGGCCAACCTCTGGGGGGCATG	2040
Qy	901	GGAGTGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	960	Db	3577	CACCATCTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	3636
Db	2497	GGAGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	2556	Qy	2041	GGC 2043	
Qy	961	GACCTGGTCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020	Db	3637	GGC 3639	
Db	2557	GACCTGGTCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2616				
Qy	1021	TGGCAGGGAGAACTGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1080				
Db	2617	TGTGCAACAACTACTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	2676				
Qy	1081	CGCGTGTAGGCTCTGGCTCTGGGGGGTAACAGTGTCTCCGGGAA	1140				
Db	2677	CGCGTGTAGGCTCTGGCTCTGGGGGGGGGGGGGGGGGGGGGGGG	2736				
Qy	1141	AGGACGCCGGCAGCACCGTCACTCAGTCCTCCGGGGGGGGGGGG	1200				
Db	2737	AGGACGCCGGCAGCACCGTCACTCAGTCCTCCGGGGGGGGGGGG	2796				
Qy	1201	AGGAGGCTCCACAGTGGATTAACGAGGACTCTCCACGCCCTCGGCTG	1260				
		RESULT 4					
		ACA6111					
		ID ACA6111 standard; cDNA; 5965 bp.					
		XX					
		ACA6111;					
		AC					
		DT 09-JUL-2003 (first entry)					
		XX					
		DE HCV cDNA encoding Met-NS3-NS4B-NS5A-NS5B (inactivated).					
		XX					
		HCV; 5'; gene; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B; KW; adenoviral vector; HCV infection; vaccine; mutant; gene therapy; protease.					

OS	Hepatitis C virus.
synthetic.	
FH	location/Qualifiers
FT	7..594
FT	/tag= a
FT	/product= "Met-NS3-NS4A-NS4B-NS5A-NS5B (inactive)"
FT	replace(5137..5145,GGAGACGC)
FT	/*tag= b
FT	/note= "Mutation produces inactive NS5B"
XX	wo2003031588-A2.
PN	17-APR-2003.
XX	10-OCT-2002; 2002WO-US032512.
XX	11-OCT-2001; 2001US-0328655P.
PR	13-MAR-2002; 2002US-0163774P.
XX	(MERI ) MERCK & CO INC.
PA	(RICE-) ISTRICHE BIOL MOLECOLARE ANGELETTI.
XX	Emini EA, Kaslow DC, Bett AJ, Shiver JW, Nicosia A, Lahm A;
PI	Luzzago A, Cortese R, Colloca S;
XX	WPI; 2003-381708/36.
DR	PT-P-PSDB; ABU09574.
XX	New nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A-NS5B polypeptide, useful as a component of an adenovector or DNA plasmid vaccine for preventing or treating hepatitis C virus.
PS	Claim 11; Fig 2; 231pp; English.
XX	The invention relates to a nucleic acid encoding a Met-NS3-NS4A-NS4B-NS5A -NS5B (NS stands for non-structural protein from HCV) Polypeptide appearing as ABU09574. The encoded polypeptide has sufficient protease activity to process itself to produce an NS5B protein that is enzymatically inactive. Also included are a cultured recombinant cell comprising the novel nucleic acid making an adenovector (comprising: (a) cassette by homologous recombination between the novel nucleic acid and a nucleic acid comprising a first adenovirus region from base pair 1-450 corresponding to either Ad5 or Ad6, a second adenovirus region from base pair 3511-5548 corresponding to Ad5 or from base pair 3508-5541 corresponding to Ad6, joined to the first region, a third adenovirus region from base pair 5549-28133 corresponding to Ad5 or from base pair 5542-28156 corresponding to Ad6, joined to the second region, a fourth adenovirus region from base pair 30818-33966 corresponding to Ad5 or from base pair 30799-33784 corresponding to Ad6, joined to the third region, and a fifth adenovirus region from base pair 33967-35935 corresponding to Ad5 or from base pair 33785-35759 corresponding to Ad6, joined to the fourth region, and (b) rescuing the recombinant adenovirus from the recombinant adenovirus plasmid), an adenoviral vector that is produced by method above appearing as AC61113 which has a humanised version of the polynucleotide of the invention and encodes the HCV inactivated polyprotein, a recombinant nucleic acid comprising one or more Ad6 regions and a region not present in Ad6, where at least one Ad6 region is selected from E1A, E1B, E2B, E2A, E4, L1, L2, L4 and L5. The nucleic acid is useful as a component of an adenoviral vector or a DNA plasmid vaccine providing a broad range of antigens for generating an HCV-specific cell mediated immune response. The nucleic acid may also be used in treating patients infected with HCV. The present sequence is the non-humanised cDNA encoding the HCV polyprotein with an inactive NS5B protein
Db	1603 TGGAGAGTGTCTCACAGGCTCACCCACATAGATGCAACACTCTTCAGACAGCAG 1662
QY	61 CAGCCAGAGAACACTTCCTCATCTGGATCAAATGCGAAGGTCTATACTGCTGGCTAGGGC 120
Db	1663 CAGGGAGGAACATCTCCCTACTCTGGTAGATACAGGCAAGGTCAGTGCGCCAGGCT 1722
Db	121 CAGGCCACCTCCATCTGGATCAAATGCGAAGGTCTATACTGCTGGCTAGGGCTAAGCCTACT 180
QY	1723 CAGGCCACCTCCATCTGGTAGATACAGGCAAGGTCAGTGCGCCAGGCT 1782
Db	181 CTGGGGGAAACACCCCTGCTGATAGCTGGAGGCCCTCAAACGAGSTCACCTC 240
QY	1783 CTGGACGGCCACACCCCTGCTGATAGCTGGAGGCCCTCAAACGAGSTCACCTC 1842
Db	241 ACACACCCCATTACCAATTCTCATCTGGCATGCTATGTCAGCCACCTGAGCTGCTACG 300
Db	1843 ACCCACCCATACCAATTACATCTGGCATGCTAGTCGGCTACCTGGAGSTGTC 1902
QY	301 AGCACTGGTGTGTTGCTGAGGAGGCTCTGAGCTCTGGCTGGCTGTATGCTGACACA 360
Db	1903 AGCACTGGTGTGTTGCTGAGGAGGCTCTGAGCTCTGGCTGGCTGACACA 1962
QY	361 GCGACGGTGTGTTGCTGAGGAGGCTCTGAGCTCTGGCTGGCTGAGCTCTGGCTGAC 420
Db	1963 GCGAGTGTGTTGCTGAGGAGGCTCTGAGCTCTGGCTGGCTGAGCTCTGGCTGAC 420
QY	421 AGGAAGAGTCTCTACCAGAGGTCTGATGAGTGGAGAGTGTGGCTGGCTGAGCTCT 480
Db	2023 AGGGAGTTCCTCTACCAGAGGTCTGATGAGTGGAGAGTGTGGCTGGCTGAGCTCT 2082
QY	481 ATCCAGCAGGGAAATGCACTGGCTGGAGGAAATGGCTGGCTGGAGCTGGCTGGAG 540
Db	2083 ATCCAGCAGGGAAATGCACTGGCTGGAGGAAATGGCTGGCTGGAGCTGGCTGGAG 2142
QY	541 ACAGCCACCAAGCAGGGAGGCGCTGCTCCGGAGGCTGAGCTGGCTGGAGCCCT 600
Db	2143 ACAGCCACCAAGCAGGGAGGCTGCTGCTCCGGAGGCTGAGCTGGCTGGAGCCCT 2202
QY	601 GAGACCTTGCGGAAACACATGTGGAACTCATCGGGATACTGACTTCTGGAGGC 660
Db	2203 GAGACATCTGGCGGAACGACATGTGGATTCTCATCAGGGATACTGAGCTGGAGGC 2262
QY	661 TGTGCACTCTGGCTGGGATCCGGAGTGTGCACTGATGGGGTCAAGCTGGCTGG 720
Db	2263 TTATCCACTCTGGCTGGGACCCGGAAATGAGCTGGCTGGCTGGAGCTTACAGCTCTAC 2232
QY	721 ACTAGCCCTCTACCAACCCATCTACCCCTCTCTTAACATCTGGGGATGGCTGG 780
Db	2323 ACCAGCCGCTCACCAAAAGTACCTCTGGCTTAACATCTGGGGATGGCTGG 2382
QY	781 GCCCAACTCTCCCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 840
Db	2383 GCCCAACTCTGGCCCTCCCGAGCGCTCTGGCTGGCTGGCTGGCTGGCTGG 2442
QY	841 GCTGTTGGCGCATAGGCCCTGGAAAGGGCTCTGGGATCTGGGGACTATGGAGA 900
Db	2443 GCTGTTGGCGCATAGGCCCTGGAAAGGGCTCTGGGATCTGGGGACTATGGAGA 2502
QY	901 GAGAGTGGCGAGCATAGGCCCTGGAAAGGGCTCTGGGATCTGGGGACTATGGAGA 960
Db	2503 GAGAGTGGCGAGCATAGGCCCTGGAAAGGGCTCTGGGATCTGGGGACTATGGAGA 2562
QY	961 GACCTGGTGTGTTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1120
Db	2563 GACCTGGTGTGTTGCTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2622

Query Match Similarity 88.3%; Score 1803; DB 8; Length 5965;  
 Best Local Similarity 92.7%; Pred. No. 0;  
 Matches 1893; Conservative 0; Mismatches 150; Indels 0;  
 Gaps 0;

Db	2683	CGCTGATAGCGTCGCTCGCGGTTATCAGTTCCCCAGGACATGCGCTGAG	2742	DT
Qy	1141	AGCGAACCGCGAGCACGTRGACTCTAGTCTCCPCCGACCTTACTATACCCAACTGTRG	1200	XX
Db	2743	AGCGAACCGCGAGCGCGTTACTAGTCTCCPCCGACCTTACTATACCCAACTGTRG	2802	DE
Qy	1201	AGAGAGTCCACCAAGCTGGATTACAGGAGCTGTCAGGCCCCTGCGGCT	1260	XX
Db	2803	AAAAGGTTCCACCAAGCTGGATTACAGGAGCTGTCAGGCCCCTGCGGCT	2862	KW
Qy	1261	AGGGAGTGTGGGACTGGATATGCCAAGTTGGCTCAAGAGACTGCTCACACCGTGTCCGGCT	1320	HCV; ds; non-structural protein; NS3; NS4A; NS4B; NS5A; NS5B; adenoviral vector; HCV infection; vaccine; gene therapy; protease.
Db	2863	AGGGAGTGTGGGACTGGATATGCCAAGAGACTGCTCACACCGTGTCCGGCT	2922	OS
Qy	1321	AAGCTCTGCGCGGATACCGGAGTCGGTCCCTTCTCGACGTTAACAGGG	1380	OS
Db	2923	AAGCTCTGCGCGGATACCGGAGTCGGTCCCTTCTCGACGTTAACAGGG	2982	OS
Qy	1381	GCTGGGGAGGAGGGATCATGCGAGACCTCTCATGCGAGACAGATACCGGA	1440	OS
Db	2983	GCTGGGGAGGAGGGATCATGCGAGACCTCTCATGCGAGACAGATACCGGA	3042	Synthetic.
Qy	1441	CATGTCAAAACGGTTCATGAGATCGTGGCCCTAAGACCTGTAATGCGAT	1500	XX
Db	3043	CATGTCAAAACGGTTCATGAGATCGTGGCCCTAAGACCTGTAATGCGAT	3102	XX
Qy	1501	CGNACATCCCGATCACGCATCACCCACGGGCCCCTGACGCCCTCCAGCCCAAC	1560	PR
Db	3103	GGAACTTCCCATCACGCATCACCCACGGGCCCCTGACGCCCTCCAGCCCAAC	3162	PR
Qy	1561	TATTCAGGGCTGTGGGGTGGCGCTGAGGAGTACGGAGTACGGGGGGGG	1620	PR
Db	3163	TATTCAGGGCTGTGGGGTGGCGCTGAGGAGTACGGAGTACGGGGGGGG	3222	PR
Qy	1621	GATTTCACTAGTGACGAGCATGACCACTGACAACTGAAAGTAAATGCCGRCGAGTTCA	1680	PR
Db	3223	GATTTCACTAGTGACGAGCATGACCACTGACAACTGAAAGTAAATGCCGRCGAGTTCA	3282	PR
Qy	1681	GCCCCGAACTCTCACAGAAGTGGATGGGGTGGCTCACAGGTACCTCCGGGTGC	1740	PT
Db	3283	GCCCCGAACTCTCACAGAAGTGGATGGGGTGGCTCACAGGTACCTCCGGGTGC	3342	PT
Qy	1741	AAACCTCTCTAGGGAGGGTCAATTCAGGTCAACCAAATACCTGGTGG	1800	PT
Db	3343	AGGCCCTCTCTAGGGAGGGTCAATTCAGGTCAACCAAATACCTGGTGG	3402	PT
Qy	1801	TGCGACATCCATGGAGCCGACGCGATGAGCTCACTTCCATGTCACCGAC	1860	PS
Db	3403	TGCGACATCCATGGAGCCGACGCGATGAGCTCACTTCCATGTCACCGAC	3462	PS
Qy	1861	CCCTCCCATCACAGAAGAGCGGCTAAGCGAGCTGGCTGGGGTCCCCCTGC	1920	XX
Db	3463	CCCTCCCATCACAGAAGAGCGGCTAAGCGAGCTGGGGTCCCCCTGC	3522	XX
Qy	1921	TGGCAGCTCTCGAGCTAGCGTGTGCTGGCCCTCTCGAGGGCATCATTAAC	1980	CC
Db	3523	TGGCAGCTCTCGAGCTAGCGTGTGCTGGCCCTCTCGAGGGCATCATTAAC	3582	CC
Qy	1981	CAAATGACTTCCAGAGCTGACTCATCGAGGCAACCTCTGTGGGGCATGAGTC	2640	CC
Db	3583	CAAATGACTTCCAGAGCTGACTCATCGAGGCAACCTCTGTGGGGCATGAGTC	3642	CC
Qy	2041	GCG 2643		CC
Db	3643	GCG 3645		CC
RESULT 5				
ID	ACA6113	ACA6113 standard; DNA; 37090 BP.		
XX				
AC	ACA6113;			

	Sequence	37090	BP;	8382	A;	10762	C;	10299	G;	7647	T;	0	U;	0	Other;																																																																																		
	Query	1	TCGGAAAGGCGCTTCAAGGCCTCACCCACGGTGTGATGCCACTCTTGCCCAAACAAAG	2860	TGGGAGAGTGTCTCAAGGCCCTCACCCACATAGATGACACTCTGTGTCAGACAGCAAG	61	CGGGCAGAGCAACTTCCCCTACCTGTGCGTGTACAGGGTACTCTGTGCGTGGCC	QY	2920	CAGGAGAGAACACTTCCCCTACCTGTGCGTGTACAGGGTACTCTGTGCGTGGCC	QY	121	CAGGCCAACCTCATATGGGATCAATGTGAAGTGTCTCATAGGCTAAAGCTACT	Db	2980	CAGGCCAACCTCATATGGGATCAATGTGAAGTGTCTCATAGGCTAAAGCTACT	QY	181	CTGGCGGGCCACACCCCTTGTATAAGGTGTGGAGGCCGTCACAAACAGGGTACACCTC	Db	3040	CTGCAAGGCCACACCCCTTGTGTACAGGTGTGGAGGCCGTCACAAATGAGGTACCCCTC	QY	241	ACACACCCATAAACCAATTCACTACGACATCATTCAGGACACTGGGTCTCAGC	Db	3100	ACCCACCCATAAACCAATTCACTACGACATCATTCAGGACACTGGGTCTCAGC	QY	301	AGCACCTGGTGTGGCGGGGTCCTTGAGCTGTGGCTGGGATTTGCTGAGACA	Db	3160	AGCACCTGGTGTGGCGGGGTCCTTGAGCTGTGGCTGGGATTTGCTGAGACA	QY	361	GGCAGCGGTGTCATTGGTAGGATCATCTGTGCGGGGCCGCTATATGTTCCGAC	Db	3220	GGCAGTGTGTGTCATTGGTAGGATCATCTGTGCGGGGCCGCTATATGTTCCGAC	QY	421	AGGGAGTCTTACCAAGGAGTGTGAGTGAATGGAGAGTGGCTGGGACTCTCC	Db	3280	AGGGAGTCTTACCAAGGAGTGTGAGTGAATGGAGAGTGGCTGGGACTCTCC	QY	481	ATGGGAGGATGAGCTCCAGGAGCAGTCAGAAGAAAAGCGTGGGGTGTGAG	Db	3340	ATGGGAGGATGAGCTCCAGGAGCAGTCAGAAGAAAAGCGTGGGGTGTGAG	QY	541	ACAGCACCAGAACAGGGAGGCCGCGCTCCCGTGTGGACTCCAGTGGGAGSCCT	Db	3400	ACAGCACCAGAACAGGGAGGCCGCGCTCCCGTGTGGACTCCAGTGGGAGSCCT	QY	601	GAGACCTCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	Db	3460	GAGACATCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	661	TGGTCACTCTSGGAAATCCGGATGATGATCACTGATGGGTTACAGGCTTGTC	Db	3520	TGTTACACTCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	721	ACTGGCCCTTCAACCCATCTACCTCCCTCTACATCTGGGGATGGGAGCC	Db	3580	ACAGGCGCTCACCAACCAAGTAACCTCTCTTACATCTGGGGGTGGGTGCT	QY	781	GCCCAACTCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	Db	3640	GCCCCACTGCCCCCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	841	GCTGTGTCAGCATGGCTTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	Db	3700	GCTGTGTCAGCATGGCTTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	901	GGAGTGGGAGGGCGCTCTGGCTTAAAGTGTGACGGCGCAATCCCTCCACCGAG	Db	3760	GGAGTGGGAGGGCGCTCTGGCTTAAAGTGTGACGGCGCAATCCCTCCACCGAG	QY	961	GACCTGGTTAACTTACTCTCTGCTATCCCTCTCTGCTGCTGCTGCTGCTGCTG
CC	Sequence	37090	BP;	8382	A;	10762	C;	10299	G;	7647	T;	0	U;	0	Other;																																																																																		
XX	Query	1	TCGGAAAGGCGCTTCAAGGCCTCACCCACGGTGTGATGCCACTCTTGCCCAAACAAAG	2860	TGGGAGAGTGTCTCAAGGCCCTCACCCACATAGATGACACTCTGTGTCAGACAGCAAG	61	CGGGCAGAGCAACTTCCCCTACCTGTGCGTGTACAGGGTACTCTGTGCGTGGCC	QY	2920	CAGGAGAGAACACTTCCCCTACCTGTGCGTGTACAGGGTACTCTGTGCGTGGCC	QY	121	CAGGCCAACCTCATATGGGATCAATGTGAAGTGTCTCATAGGCTAAAGCTACT	Db	2980	CAGGCCAACCTCATATGGGATCAATGTGAAGTGTCTCATAGGCTAAAGCTACT	QY	181	CTGGCGGGCCACACCCCTTGTATAAGGTGTGGAGGCCGTCACAAACAGGGTACACCTC	Db	3040	CTGCAAGGCCACACCCCTTGTGTACAGGTGTGGAGGCCGTCACAAATGAGGTACCCCTC	QY	241	ACACACCCATAAACCAATTCACTACGACATCATTCAGGACACTGGGTCTCAGC	Db	3100	ACCCACCCATAAACCAATTCACTACGACATCATTCAGGACACTGGGTCTCAGC	QY	301	AGCACCTGGTGTGGCGGGTCCTTGAGCTGTGGCTGGGATTTGCTGAGACA	Db	3160	AGCACCTGGTGTGGCGGGTCCTTGAGCTGTGGCTGGGATTTGCTGAGACA	QY	361	GGCAGCGGTGTCATTGGTAGGATCATCTGTGCGGGGCCGCTATATGTTCCGAC	Db	3220	GGCAGTGTGTGTCATTGGTAGGATCATCTGTGCGGGGCCGCTATATGTTCCGAC	QY	421	AGGGAGTCTTACCAAGGAGTGTGAGTGAATGGAGAGTGGCTGGGACTCTCC	Db	3280	AGGGAGTCTTACCAAGGAGTGTGAGTGAATGGAGAGTGGCTGGGACTCTCC	QY	481	ATGGGAGGATGAGCTCCAGGAGCAGTCAGAAGAAAAGCGTGGGGTGTGAG	Db	3340	ATGGGAGGATGAGCTCCAGGAGCAGTCAGAAGAAAAGCGTGGGGTGTGAG	QY	541	ACAGCACCAGAACAGGGAGGCCGCGCTCCCGTGTGGACTCCAGTGGGAGSCCT	Db	3400	ACAGCACCAGAACAGGGAGGCCGCGCTCCCGTGTGGACTCCAGTGGGAGSCCT	QY	601	GAGACCTCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	Db	3460	GAGACATCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	661	TGGTCACTCTSGGAAATCCGGATGATGATCACTGATGGGTTACAGGCTTGTC	Db	3520	TGTTACACTCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	721	ACTGGCCCTTCAACCCATCTACCTCCCTCTACATCTGGGGATGGGAGCC	Db	3580	ACAGGCGCTCACCAACCAAGTAACCTCTCTTACATCTGGGGGTGGGTGCT	QY	781	GCCCAACTCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	Db	3640	GCCCCACTGCCCCCTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	841	GCTGTGTCAGCATGGCTTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	Db	3700	GCTGTGTCAGCATGGCTTGGGGAAACAGTGGGACTTCATGGGGATAAGTATTAGGAGC	QY	901	GGAGTGGGAGGGCGCTCTGGCTTAAAGTGTGACGGCGCAATCCCTCCACCGAG	Db	3760	GGAGTGGGAGGGCGCTCTGGCTTAAAGTGTGACGGCGCAATCCCTCCACCGAG	QY	961	GACCTGGTTAACTTACTCTCTGCTATCCCTCTCTGCTGCTGCTGCTGCTG
Best Local Similarity	88.3%	Score	1803	DB	8	Length	37090																																																																																										
Matches	1893	Conservative	92.7%	Pred.	No.	0	Mismatches	150	Indels	0	Gaps	0																																																																																					

Db	3820	GACCTGGTCAATCCTACTTCCTGCACTCCCTCTCTCTGCGCTTGCGCTCGCGGTCCTG
Qy	1021	TGCCAGGCAATACTCGCGTGGCAGCTGGTCCAGGGGGGSSCTGTGCACTGATGATGAC
Db	3880	TGTGCGAATACTCGTGTGCGATGAGCTGGTCCGGAGGGGGCTGTGAG
Qy	1081	CGGTGATACCGTRGCTCGCGGGTAACTATTTCCCCCGCATATCGGCCAG
Db	3940	CGGCGATAGCTCGCGGGTAACTATCGTGTGCGCTCGGCCAGGGGGCTGTGAG
Qy	1141	AGCAGGCCAGCGAGCTGCACTCAGATCTCTCGACGACTATGCTCGACGACTATGCG
Db	4000	AGCGGCCAGCGCGTGTACCTCGATCTCCACCGCTTACCATCGTGTG
Qy	1201	AAGGGCTCACCGCTGCACTCAGAGACTGCTCCAGGCCCTCGCGGCTGTGCT
Db	4060	AAAAGCTCCACCGTGTGAACTAATGAGACTGCTCCACACCGTGTGCGCTGTGGCT
Qy	1261	AGGAGTTGGAGCTGGATATGCACTGACAGTTGCTGACTCGAGACTGGCTCAG
Db	4120	AGGAGTTGGAGCTGGATATGCACTGACAGTTGCTGACTCGAGACTGGCTCAG
Qy	1321	AAGCTCTCGCGCAATTACCGGGAGTCCTCTTCTCATGCCCACCGTGGTACAAAGGG
Db	4180	AGCTCTCGCGCACTATCGGGATCTCCCTTCTCATGCCCACCGTGGTACAAAGGG
Qy	1381	GTCTGGCGGAGCGGGCTCATCGAGCACCTGCTATGAGAGCACAGATACCGGA
Db	4240	GTCTGGCGGAGAGCGGCATCATCGAACACCTGCTCCACCGTGGTACAAAGGG
Qy	1441	CATGCTAAACAGGTTCCATGAGGAATCTGGGCTTAAGACCTGAGTACATGTTGGAT
Db	4300	CATGCTAAACAGGTTCCATGAGGAATCTGGGCTTAAGACCTGAGTACATGTTGGAT
Qy	1501	GGAACTATCCCGATCAACCATAACCGGCGGCTGACGCCCTCCAGGCCAAC
Db	4360	GGAACTATCCCGATCAACCGGCGGCTGACGCCCTCCAGGCCAAC
Qy	1561	TATTCAGGGCGCTGGGGTGCTGCTGAGGAGTCTGGGGTAGCGGGGTGGGG
Db	4420	TATTCAGGGCGCTGGGGTGCCSCTGAGGAGTCTGGGGTAGCGGGGTGGGG
Qy	1621	GATTTCACTACAGGCACTAACCACTGACAGCTAACCTGACAGTAAATGCCAGGTTCA
Db	4480	GATTTCACTACAGGCACTAACCACTGACAGCTAACCTGACAGTAAATGCCAGGTTCA
Qy	1681	GCCCCCGAAATCTTCAAGAGGTACAGGAGTGTGGGTGGGGCTCACAGTGCTCCGGTGC
Db	4540	GCTCCGAAATCTTCAAGGGGTACAGGAGTGTGGGTGGGGCTCACAGTGCTCCGGTGC
Qy	1741	AAA CCTCTCTACCGGAGGGTACATCCAGGTCGGCTCACAGTGCTCCGGTGC
Db	4600	AAA CCTCTCTACCGGAGGGTACATCCAGGTCGGCTCACAGTGCTCCGGTGC
Qy	1801	TCGCGCTCCATCGAGCGCAACCGGAGTACAGGAGTGTGGGTGGGGCTCAC
Db	4660	TCAGCTTACATCGAGCGCAACCGGAGTGTGGGTGGGGCTCAC
Qy	1861	CCCTCCACATCACCGAGAGACCGCTAAGGCGCTGGGGGCTCCCTGAGGCACTAC
Db	4720	CCCTCCACATCACCGAGAGACCGCTAAGGCGCTGGGGGCTCCCTGAGGCACTAC
Qy	1921	TTGCCAGCTTCCCGAGCTGAGCTCATCGAGGCCAACCTCCCTGGGGCATAGAGT
Db	4780	TTGCCAGCTTCCCGAGCTGAGCTCATCGAGGCCAACCTCCCTGGGGCATAGAGT
Qy	1981	CAAATGACTTCCCGAGCTGAGCTCATCGAGGCCAACCTCCCTGGGGCATAGAGT
Db	4840	CAACATGCTCTCCCGAGCGCTGAGCTCATCGAGGCCAACCTCCCTGGGGCATAGAGT
Qy	2041	GGC 2043

Db 4900 GGC 4902  
 CC compounds against different HCV isolates and facilitating HCV research.  
 CC where compounds that inhibit HCV replication have research and  
 CC therapeutic applications in identifying overall efficacy and lack of  
 CC unacceptable toxicity. Accordingly, they can be used to treat or inhibit  
 CC the onset of HCV in a patient. This polynucleotide is a mutant HCV  
 CC replicon DNA sequence that encodes an HCV NS1-NSSB Polyprotein (based on  
 CC HCV-BK) with 41M and 1179I residue substitutions, given in an  
 CC exemplification of the invention.

XX sequence 7987 BP; 1722 A; 2315 C; 2204 G; 1746 T; 0 U; 0 Other;  
 XX 18-NOV-2004 (first entry)  
 DE DNA encoding Hepatitis C virus (isolate BK) NS3-NSSB polyprotein Seq 3 .  
 KW Hepatitis C virus; HCV; NS3; NSSA; non-structural protein;  
 KK virus replication activity; gene; ds; mutant.  
 OS Hepatitis C virus (isolate BK).  
 OS Synthetic.

XX  
 FH Key Location/Qualifiers  
 FT 5'UTR 1. 386  
 FT /tag= a  
 FT 387. .1181  
 FT /tag= b  
 FT misc\_feature /product= "Beta-lactamase"  
 FT 1225. .1800  
 FT /tag= c  
 FT /label= EMCV IRES  
 FT /note= "internal ribosome entry site"  
 FT 1801. .7758  
 FT /\*tag= d  
 FT /product= "NS3-NS5B polyprotein"  
 FT /gene= "NS3 (1804-696)"  
 FT /gene= "NS4A (3697-3851)"  
 FT /gene= "NS4B (3859-4641)"  
 FT /gene= "NS5A (4642-5982)"  
 FT /gene= "NS5B (5983-7755)"  
 FT 7759. .7987  
 FT /\*tag= e

XX WO2004074507-A2.  
 PD 02-SEP-2004.  
 XX 09-FEB-2004; 2004WO-US003726.  
 PR 13-FEB-2003; 2003US-0447318P.  
 PA (MERI ) MERCK & CO INC.  
 XX Grobler J, Flores O, Markel BJ;  
 PI WPI; 2004 635590/61.  
 DR P-PSDB; ADR38450.  
 XX  
 PT Making Hepatitis C virus (HCV) replicon having increased replication  
 PT activity, useful in HCV research, comprises modifying HCV replicon  
 PT construct to encode an amino acid substitution at a position  
 PT corresponding to amino acid 470 of NS3.  
 XX  
 PS Claim 10, SEQ ID NO 3; 54pp; English.

XX  
 CC This invention relates to a novel method for producing a Hepatitis C  
 CC virus (HCV) replicon having an increased replication activity.  
 CC Specifically, it refers to modifying an HCV replicon construct to encode  
 CC an amino acid substitution in NS3 (a non-structural protein that along  
 CC with NS4A, NS4B, NS5A and NS5B make up the virus replication machinery  
 CC released in the form of a polyprotein). The present invention describes  
 CC an amino acid substitution at a position corresponding to amino acid 470  
 CC of NS3 alone, or in combination with, an isoleucine in a position  
 CC corresponding to amino acid 232 of NS5A that confers improved cell  
 CC replication activity compared to wild type HCV. The method is  
 CC useful for facilitating the identification of broadly efficacious

Query Match 88.2%; Score 1801.4; DB 13; Length 7987;  
 Best Local Similarity 92.6%; Pred. No. 0; Mismatches 151; Indels 0; Gaps 0;  
 Matches 1892; Conservative 0; Mi-matches 3397; Score 1801.4; DB 13; Length 7987;  
 Qy 1 TCGGAGGGCGGTTCAAGGCTCACCCACCTGGATGCCACTTCTGTGCTCAAACAAAG 60  
 Db 3457 CAGGAGGAGACAATTCCTAACCTGCTAGTGATCACCAAGCAGCTGCGCCAGGGT 3516  
 Qy 121 CAGGCCAACCTCATATGGGATCAATGTTGAAGTGTCATAGCGCTAAAGCTACT 180  
 Db 3517 CAGGCCAACCTCATATGGGATCAATGTTGAAGTGTCATAGCGCTAAAGCTACT 3456  
 Qy 181 CTGGGGGGCCACACACCTGCTGTATAGGTGGAGCCGTCAMACAGAGTCACCTC 240  
 Db 3577 CTCGACGGCCACACCTGCTGTAGGCTGACAGCTGGAGGCGCTAACATGAGTCACCTC 3636  
 Qy 241 AACACACCCATTACCAATTATCATCATGCTCATGTCAGCGACCTGGAGGTGTCACG 300  
 Db 3637 ACCACACCCATTACCAATTATCATCATGCTCATGTCAGCGACCTGGAGGTGTCACG 3696  
 Qy 301 AGCACCTGGTCTGGGGGGGCTCTGGAGCTGGCTGGCTGGCTGGTGTGACACACA 360  
 Db 3697 AGCACCTGGTCTGGGGGGGCTCTGGAGCTGGCTGGCTGGCTGGTGTGACACACA 3756  
 Qy 361 GCGAGCTGGTATGGGTAGGATCATCTGTCGGGGGGCCGCTATGTTCCCGAC 420  
 Db 3757 GCGAGCTGGTATGGGTAGGATCATCTGTCGGGGGGCCGCTATGTTCCCGAC 3816  
 Qy 421 AGGGAGCTCTTACAGGGAGTCGATGAGATGGAGAGGGCTGGCTGGCACCCTCTAC 480  
 Db 3817 AGGGAGCTCTTACAGGGAGTCGATGAGATGGAGAGGGCTGGCCTGGCACCCTCTAC 3876  
 Qy 481 ATCGAGAGGGATGCGCTGCCGAGCAGTCAGCAACAAAAGCGCTCGGGTGCAGCAG 540  
 Db 3877 ATCGAGAGGGATGCGCTGCCGAGCAGTCAGCAACAAAAGCGCTCGGGTGCAGCAG 3936  
 Qy 541 ACAGCCACCAAGCAAGGGAGCCGGTCTCGCTCGCTGGGGATGGTCAAGTGGAGGCCCT 600  
 Db 3937 ACAGCCACCAAGCAAGGGAGCCGGTCTCGCTGGGGATGGTCAAGTGGAGGCCCT 3996  
 Qy 601 GAGACCTCTGGGGAAACATGTGGACTCATGGGGATAAGTACTTAGGGC 660  
 Db 3997 GAGACATCTGGGGAGACAGCTGCGTGAATTCTAGGGGGATGGTACAGTC 4056  
 Qy 661 TGTGCTCTCTGGGGATTGATCATCTGCTGGGGTACAGCTCTGTC 720  
 Db 4057 TATCCACTCTGGGGACCCGAGATGATCTGATGGCATACGCCCTATC 4116  
 Qy 721 ACTAGCCGCTACCCGCACTAACCTCTCCCTGCTAACATCTGGGGATGGAGCC 780  
 Db 4117 ACCAGCCGCTACCCGCACTAACCTCTCCCTGCTAACATCTGGGGATGGAGCC 4176  
 Qy 781 GCGCAACTCGCTCCCCCAGTCGCTGCTGCTGAGGGGGCATCTGTGCG 840  
 Db 4177 GCGCAACTCGCCCCCAGGGCCGCTGCTGCTGAGGGGGCATCTGTGCG 4236  
 Qy 841 GCGTGTGGCAGCATGGCCCTGGAGGTGGCTGAGCACATCTGGGGCTATGGAGCA 900

Db	4237	GCTGTGTTGGCACATAGCCTTGGGAAGGTGTTGACATTCCTGGGTTATGGCA	4296	Qy	1981	CAAATGACTCCAGAGCGTGACCTCATGGGCCAACCTCTCTGGCGCATGAGTG	2040
Qy	901	GGAGTGGCCAGGGCGCTGTGTCCTCAAGGTGACGGCAGAATGCGCTCACCGAG	960	Db	5377	CACCATCTCTCCGGAGCTGACCTCATGGGCCAACCTCTGGCGCAGAGATG	5436
Db	4297	GGAGTGGCCAGGGCGCTGTGTCCTCAAGGTGACGGCAGAATGCGCTCACCGAG	4356	Qy	2041	GCC 2043	
Qy	961	GGCTGGTTAACTTACTCTCCATCTCTCTGGCTGCGGCTGCGGTGCGGTG	1020	Db	5437	GACCTGGTCATCTACTCTCTCCATCTCTCTGGCTGCGGCTGCGGTG	4416
Db	4357	GACCTGGTCATCTACTCTCTCCATCTCTCTGGCTGCGGCTGCGGTG	4416	Qy	1021	TGGCGAGGAGATACTGGTCGCGACGCGGGTCAAGGGAGGGCTGAGTGGTGA	1080
Qy	4417	TGTGCAAGCAATTACTGGCTGACAGCGTGGTGGTCGGAGAGGGGGCTGAGTGGTGA	1080	Db	4417	TGTGCAAGCAATTACTGGCTGACAGCGTGGTGGTCGGAGAGGGGGCTGAGTGGTGA	4416
Db	1081	CGCGTGTAGCTCGCTCGGGGTACCGAGTGTCCAGTGTCCCACCGACTATGTGCGAGAG	1140	Qy	1081	CGCGTGTAGCTCGCTCGGGGTACCGAGTGTCCAGTGTCCCACCGACTATGTGCGAGAG	1140
Db	4477	CGCGTGTAGCTGGTGCCTCGGGGTAAATAGTTCCACAGCTGCTGCG	4536	Db	4477	CGCGTGTAGCTGGTGCCTCGGGGTAAATAGTTCCACAGCTGCTGCG	4536
Qy	1141	AGCGAGGAGCCGAGGAGCAGCTGTCAGTCTCGAGCTTCTCGAGCTTACTATCACCAACTGTG	1200	Qy	1141	AGCGAGGAGCCGAGGAGCAGCTGTCAGTCTCGAGCTTACTATCACCAACTGTG	1200
Db	4537	ACCGACGCCGAGCGCTGCTACTCAGATCTCCACCTTACATCAGCTGCTG	4596	Db	4537	ACCGACGCCGAGCGCTGCTACTCAGATCTCCACCTTACATCAGCTGCTG	4596
Qy	1201	AGAGGTCACCAGGGATTACAGGAGCTCTCCACGCCCTGCTCGGCTGCGCTA	1260	Db	4597	AAAAGCTCCACCAAGGTGATTAATGAGACTCTCCACGCCCTGCTCGGCTA	4656
Db	4597	AAAAGCTCCACCAAGGTGATTAATGAGACTCTCCACGCCCTGCTCGGCTA	4656	Qy	1261	AGGATGTTGGACTGGATGAGACTCTCCACGCCCTGCTCGGCTA	1320
Qy	1261	AGGATGTTGGACTGGATGAGACTCTCCACGCCCTGCTCGGCTA	1320	Db	4657	AGGAGTTGGACTGGATGAGACTCTCCACGCCCTGCTCGGCTA	4716
Db	4657	AGGAGTTGGACTGGATGAGACTCTCCACGCCCTGCTCGGCTA	4716	Qy	1321	AGCTCTGGCGGATPACCGAGGTCCTTTCTCATCCAAGCTGAGGAAAGGG	1380
Qy	1321	AGCTCTGGCGGATPACCGAGGTCCTTTCTCATCCAAGCTGAGGAAAGGG	1380	Db	4717	AGCTCTGGCGGATPACCGAGGTCCTTTCTCATCCAAGCTGAGGAAAGGG	4776
Db	4717	AGCTCTGGCGGATPACCGAGGTCCTTTCTCATCCAAGCTGAGGAAAGGG	4776	Qy	1381	GTCTGGGGAGACGCAATATGAGACGCCCTGCTGAGACAGATCACCGA	1440
Qy	1381	GTCTGGGGAGACGCAATATGAGACGCCCTGCTGAGACAGATCACCGA	1440	Db	4777	GTCTGGGGAGACGCAATATGAGACGCCCTGCTGAGACAGATCACCGA	4836
Db	4777	GTCTGGGGAGACGCAATATGAGACGCCCTGCTGAGACAGATCACCGA	4836	Qy	1441	CATGTCAAAAGGGTTCATGGATGTTGGCTTAAGACCTGTTAGTGGCAT	1500
Qy	1441	CATGTCAAAAGGGTTCATGGATGTTGGCTTAAGACCTGTTAGTGGCAT	1500	Db	4837	CATGTCAAAAGGGTTCATGGATGTTGGCTTAAGACCTGTTAGTGGCAT	4896
Db	4837	CATGTCAAAAGGGTTCATGGATGTTGGCTTAAGACCTGTTAGTGGCAT	4896	Qy	1501	GGACATTCCTCATCACGCAACACCGGCCCCCTGAGCCCTCCGGCCAAC	1560
Qy	1501	GGACATTCCTCATCACGCAACACCGGCCCCCTGAGCCCTCCGGCCAAC	1560	Db	4897	GGACATTCCTCATCACGCAACACCGGCCCCCTGAGCCCTCCGGCCAAC	4956
Db	4897	GGACATTCCTCATCACGCAACACCGGCCCCCTGAGCCCTCCGGCCAAC	4956	Qy	1561	TATTCAGGGCTGTGGCTGGCTGGCTGGAGCTACGTGAGGTTACCGGGGGGG	1620
Qy	1561	TATTCAGGGCTGTGGCTGGCTGGCTGGAGCTACGTGAGGTTACCGGGGGGG	1620	Db	4957	TATTCAGGGCTGTGGCTGGCTGGAGCTACGTGAGGTTACCGGGGGGG	5016
Db	4957	TATTCAGGGCTGTGGCTGGCTGGAGCTACGTGAGGTTACCGGGGGGG	5016	Qy	1621	GATTCCACTACTGTGAGGAGGACTGACACCTAAAGCCCGTGCAGGTCCA	1680
Qy	1621	GATTCCACTACTGTGAGGAGGACTGACACCTAAAGCCCGTGCAGGTCCA	1680	Db	5017	GATTCCACTACTGTGAGGAGGACTGACACCTAAAGCCCGTGCAGGTCCA	5076
Db	5017	GATTCCACTACTGTGAGGAGGACTGACACCTAAAGCCCGTGCAGGTCCA	5076	Qy	1681	GGCCCGGATCTTCAGAGACTGGTGGGGGGGGGGGGGGGGGGGGGGGGGG	1740
Qy	1681	GGCCCGGATCTTCAGAGACTGGTGGGGGGGGGGGGGGGGGGGGGGGGGG	1740	Db	5077	GTCTGTGATTCCTCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5136
Db	5077	GTCTGTGATTCCTCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5136	Qy	1741	AAACCTCTCTACGGAGGGGTACATTCAGGTGGCTGCAACCAATACTGGTGG	1800
Qy	1741	AAACCTCTCTACGGAGGGGTACATTCAGGTGGCTGCAACCAATACTGGTGG	1800	Db	5137	AAACCTCTCTACGGAGGGGTACATTCAGGTGGCTGCAACCAATACTGGTGG	5196
Db	5137	AAACCTCTCTACGGAGGGGTACATTCAGGTGGCTGCAACCAATACTGGTGG	5196	Qy	1801	TGGCACTCCATGGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	1860
Qy	1801	TGGCACTCCATGGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	1860	Db	5197	TGAGCTTACGCTTACGGAGCCGACCGATGAGCTGCTACTCCATGTCACCGAC	5256
Db	5197	TGAGCTTACGCTTACGGAGCCGACCGATGAGCTGCTACTCCATGTCACCGAC	5256	Qy	1861	CCCTCCACATCACGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	1920
Qy	1861	CCCTCCACATCACGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	1920	Db	5257	CCCTCCACATCACGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	5316
Db	5257	CCCTCCACATCACGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	5316	Qy	1921	TGGCACTCCATGGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	1980
Qy	1921	TGGCACTCCATGGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	1980	Db	5317	TGGCACTCCATGGAGGGCTGACGGATGAGCTGCTACTCCATGTCACCGAC	5376

The invention describes a RNA interference (RNA) agent (I) comprising a sense sequence and an antisense sequence, where the sense sequences have one or more asymmetrical 2'-O alkyl modifications, the antisense sequences have one or more asymmetrical phosphorothioate modifications and the antisense targets a human gene sequence. Also described are: a pharmaceutical preparation comprising (I); reducing (M1) apoB-100 levels or glucose-6-phosphatase levels in a subject; producing (I);

Example 5; SEQ ID NO 6688; 378pp; English.

(ALNY-) ALNYLAM PHARM.

PA

XX

PI

XX

Manoharan M, Blumrot D;

WPI; 2004-67736/66.

DR

XX

PT

Interference RNA agent useful for treating dyslipidemias, coronary artery disease, diabetes, cancer or neurological disease, comprises sense sequence and antisense sequence which has specific modifications.

XX

PS

XX

CC

sense sequence and an antisense sequence, where the sense sequences have one or more asymmetrical 2'-O alkyl modifications, the antisense sequences have one or more asymmetrical phosphorothioate modifications and the antisense targets a human gene sequence. Also described are: a pharmaceutical preparation comprising (I); reducing (M1) apoB-100 levels or glucose-6-phosphatase levels in a subject; producing (I);







FT CDS 3351 .5177 /product= "NS2"  
 FT /tag= f  
 FT /product= "NS3"  
 FT 5178. .5.918 /tag= g  
 FT /product= "NS4a"  
 FT 5919. .6.371 /\*tag= h  
 FT /product= "NS4b"  
 FT 6372. .9.365 /\*tag= i  
 FT /product= "NS5"  
 XX EP464287-A.  
 PN PD 08-JAN-1992.  
 XX PR 28-DEC-1990; 90EP-00314371.  
 XX PR 25-JUN-1990; 90JP-00167466.  
 PR 31-AUG-1990; 90JP-00230921.  
 PR 09-NOV-1990; 90JP-00305605.  
 PR 17-JUN-1991; 91EP-00401604.  
 PA (OSAU ) UNIV OSAKA.  
 XX WPI; 1992-009617/02.  
 DR P-PSDB; AAR20091.  
 XX PT New DNA from non-A, non-B hepatitis virus - and derived antigenic  
 PT polypeptide(s) useful for diagnostics, blood screening and in vaccines.  
 PS Claim 1; FIG 2; 89pp; English.  
 XX The sequence was obtnd. from several overlapping "BK" cDNA clones obtnd. by  
 CC "gene walking" using a cDNA clone isolated from a library prep. from  
 CC NANBV RNA. The DNA and fragments of it can be used for the detection of  
 CC the presence of NANBV by hybridisation or PCR. Antigenic polypeptides  
 CC encoded by the sequence can be used as immunoassay reagents, for  
 CC screening donated blood, and as immuno- genes for vaccine prodn.  
 CC Antibodies raised to the peptides can be used in immunoassays to detect  
 CC or quantify NANBV antigens in liver tissue and blood. Preferred  
 CC polypeptides are encoded by the following nucleotides: 333-422, -677, or  
 CC -671; 474-563, -768-905; 906-953 or -1499; -1121; 1194-1322;  
 CC 1209-1322; 1500- 2519; 2520-3350; 3351-5177; 4185-4574; 5178-5919; 5544-  
 CC 5633; 5919- 6371; and 6372-9372. The sequence is also disclosed in EP-  
 CC 463848 in which a virus particle contg. antigens encoded by the sequence  
 CC is claimed, as well as expression vectors contg. the sequence. See  
 CC AAQ20268 for details of this specification. (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX Sequence 9416 BP; 1905 A; 2825 C; 2679 G; 2007 T; 0 U; 0 Other;  
 SQ Best Local Similarity 92.5%; Score 1796.6; DB 2; Length 9416;  
 Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;  
 Oy 1 TGCGAGGAGCTCTTACAGGCCCTCACCCCTGTGCTGCAACTCTCTGTCACAAAG 60  
 Db 5004 TGGAGAGAGTGCTCCACAGCTTACCATATAGTCACACTCTTGTCCGACAGCAG 5063  
 Oy 61 CAGCCGCCAACCTCCATCATGGATAAATGTGGAGATGTCATAGCGCTAAAGCTACT 180  
 Db 5124 CAGGCCAACCTCCATCATGGATAAATGTGGAGATGTCATAGCGCTAAAGCTACT 5183  
 181 CTGGCGGCCAACCCCTGTGCTGAGCTGGAGCGCTCCAAAGAGGTCAACCTTC 240  
 Qy 5184 CTGCACGGCCACACCTCTGTGCTGAGCTGGAGCGCTCCAAAGAGGTCAACCTTC 5243

QY 1321 AAGCTTCTGCCGGGATTAACCGGAGTCCCCTTTCATGCCAACGTGGTACAAAGGG 1380  
 FT FT CDS  
 DB 6324 AAGCTTCTGCCGGCAGCTTACTTGAGRCCTTUTTCGTCGAAGGGGGTACAAGGG 6383  
 FT FT CDS  
 QY 1381 GCTCTGGGGGAGACGGCATATGAGCACCTGCTCATGGAGACAGATCACCGA 1440  
 FT FT CDS  
 DB 6384 GCTCTGGGGGAGACGGCATATGAGCACCTGCTCATGGAGACAGATCACCGA 6443  
 FT FT CDS  
 QY 1441 CATGTCRAAAMGGTTCATAGGATGTTGGCTTAAGCTTGAGTAAATGTTGCA 1500  
 FT FT CDS  
 DB 6444 CATGTCRAAAMGGTTCATAGGATGTCGGCTTAAGAC 6503  
 FT FT CDS  
 QY 1501 GGAACATTCGCATCAACGATACACACGGGCCCCTGAGAACCCCTGCTCATGGAGACAGATCACCGA 1560  
 FT FT CDS  
 DB 6504 GGAACATTCGCATCAACGATACACACGGGCCCCTGAGAACCCCTGCTCATGGAGACAGATCACCGA 6563  
 FT FT CDS  
 QY 1561 TATTCCAGGGGCTGTGGCGCTGGCTAGAGGATACCGGGTGGGG 1620  
 FT FT CDS  
 DB 6564 TATTCCAGGGGCTGTGGCGCTGGCTAGAGGATACCGGGTGGGG 6623  
 FT FT CDS  
 QY 1621 GATTTCACCTAGTGGAGGACTGAGCTGGCTAGAGGATACCGGGTGGGG 1680  
 FT FT CDS  
 DB 6624 GATTTCACCTAGTGGAGGACTGAGCTGGCTAGAGGATACCGGGTGGGG 6683  
 FT FT CDS  
 QY 1681 GCCCCCGAATCTTCACTACAGAGTGGATGGGTTGGCGCTGAGCTGGAGGATACCGGGTGGGG 1740  
 FT FT CDS  
 DB 6684 GCTCTCTGAAATCTTCTGGAGGTGGAGGATACCGGGTGGGG 6743  
 FT FT CDS  
 QY 1741 AACCTCTCTTCACTACAGAGGAGGAGGATACCTCCAGGTGGCTAACCAATACATCTGGTGG 1800  
 FT FT CDS  
 DB 6744 AGGCTCTCTCCTACAGGGAGGAGGATACCTCCAGGTGGCTAACCAATACATCTGGTGG 6803  
 FT FT CDS  
 QY 1801 TCGCAGCTCCATGGAGCCGAAACGGATAGCAGTGTCACTTCATCTCACCGAC 1860  
 FT FT CDS  
 DB 6804 TCAAGCTTACCATGCGAGCCGAAACGGATAGCAGTGTCACTTCATCTCACCGAC 6863  
 FT FT CDS  
 QY 1861 CCCTCCACATCACAGAGGAGGCTAGGAGGCTGGAGGGTCTCCCTCC 1920  
 FT FT CDS  
 DB 6864 CCCTCCACATCACAGAGGAGGCTAGGAGGCTGGAGGGTCTCCCTCC 6923  
 FT FT CDS  
 QY 1921 TTGGCCAGCTTTCAGTAGGAGTGTCTGGCTTCTGAGGGACATACATTAC 1980  
 FT FT CDS  
 DB 6924 TTGGCCAGCTTTCAGTAGGAGTGTCTGGCTTCTGAGGGACATACATTAC 6983  
 FT FT CDS  
 QY 1981 CAAATGACTTCCAGAGCTGACCTCATCGAGCCACCTGTGGCGCATGAGATG 2040  
 FT FT CDS  
 DB 6984 CACCATGTCCTCGGAGCTAACCTCATCGAGGCCAACCTGTGGCGAGGATG 7043  
 FT FT CDS  
 QY 2041 GGC 2043  
 FT FT CDS  
 DB 7044 GGC 7046  
 FT FT CDS  
 RESULT 10  
 AAX53394  
 ID AAX53394 Standard; cDNA; 9416 BP.  
 XX  
 AC AAX53394;  
 XX  
 DT 20-MAR-2003 (revised)  
 DT 27-SEP-1999 (first entry)  
 Non-A, non-B hepatitis virus genome cDNA.  
 Non-A, non-B hepatitis virus: NAMEV; antigen; infection; diagnosis;  
 vaccine; db.  
 Non-A, non-B hepatitis virus  
 Non-B hepatitis virus  
 Key CDS  
 Location/Qualifiers  
 /tag= .9364  
 /tag= a  
 /tag= r  
 /tag= m  
 /tag= n  
 /product= "NS1"  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .9362  
 /\*tag= e  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .6371  
 /\*tag= d  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1499  
 /\*tag= c  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .677  
 /\*tag= b  
 /product= "Core\_protein"  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1499  
 /\*tag= a  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .905  
 /\*tag= g  
 /product= "Matrix\_protein"  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1499  
 /\*tag= f  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .474  
 /\*tag= h  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .905  
 /\*tag= i  
 /product= "Envelope\_protein"  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1499  
 /\*tag= k  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .906  
 /\*tag= j  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1121  
 /\*tag= l  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .953  
 /\*tag= m  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1046  
 /\*tag= n  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1123  
 /\*tag= o  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .1232  
 /\*tag= p  
 /product= "NS2"  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .5350  
 /\*tag= q  
 /product= "NS3"  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .4485  
 /\*tag= r  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .4574  
 /\*tag= s  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .5178  
 /\*tag= t  
 /note= "an isolated DNA comprising this sequence is specifically claimed in Claim 1"  
 333. .5918

/product= "NS4a"  
 /note= "an isolated DNA comprising this sequence is  
 specifically claimed in Claim 1"  
 5544 . 5633  
 /\*tag= S  
 /note= "an isolated DNA comprising this sequence is  
 specifically claimed in Claim 1"  
 CDS  
 5919 . 6371  
 /\*tag= t  
 /product= "NS4b"  
 /note= "an isolated DNA comprising this sequence is  
 specifically claimed in Claim 1"  
 CDS  
 6372 . 9362  
 /\*tag= u  
 /product= "NS5"  
 /note= "an isolated DNA comprising this sequence is  
 specifically claimed in Claim 1"  
 EP33426-A1.  
 XX  
 PD 04-AUG-1999.  
 XX  
 PF 28-DEC-1990; 99EP-00106005.  
 XX  
 PR 25-JUN-1990; 90EP-00167466.  
 PR 31-AUG-1990; 90EP-00230921.  
 PR 09-NOV-1990; 90EP-00305605.  
 PR 28-DEC-1990; 90EP-00314371.  
 XX  
 PA (OSAU ) UNIV OSAKA.  
 XX  
 PI Okayama H, Ruke I, Mori C, Takamizawa A, Yoshida I;  
 XX DR WPI; 1999-407152/35.  
 DR P-RSDB; AAY06423.  
 XX  
 PT New hepatitis virus polypeptides, useful for diagnosing and treating  
 hepatitis infections.  
 XX  
 PS Claim 1; Fig 2(1)-(16); 56pp; English.  
 XX  
 CC This is the nucleotide sequence of non-A, non-B hepatitis virus (NANBV)  
 CC genomic cDNA containing the entire region of the open reading frame of  
 CC the NANBV genome. To obtain the genomic cDNA, NANBV RNAs were extracted  
 CC directly from NANBV particles contained in whole blood of a patient  
 CC having NANB hepatitis, or from a resected liver of a patient having NANB  
 CC hepatitis and liver cancer. The RNA was then converted to double-stranded  
 CC cDNA. A cDNA library was produced and screened using serum from a  
 CC convalescent patient having acute NANB hepatitis and serum from a patient  
 CC having chronic NANB hepatitis. The isolated cDNA allows recombinant  
 CC production of NANBV antigen polypeptides in microbial or eukaryotic cell  
 CC culture. The method provides the safe production of NANBV antigens with  
 CC high purity on a large scale at low cost without the biohazard associated  
 CC with multiplying virus in animals. Claimed NANBV nucleotide sequences are  
 CC useful for the diagnosis of NANB hepatitis and for the recombinant  
 CC production of antigenic polypeptides. The polypeptides encoded by these  
 CC nucleotide sequences are useful as antigens for vaccines, and as  
 CC diagnostic reagents. (Updated on 20-MAR-2003 to correct PR field.)  
 (Updated on 20-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 9416 BP; 1905 A; 2826 C; 2678 G; 2007 T; 0 U; 0 Other;  
 Query Match 87.9%; Score 1796.6; DB 2; Length 9416;  
 Best Local Similarity 92.5%; Pred. No. 0;  
 Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;  
 QY 1 TGGGAGAGCTTCACAGGCCCTACCCATATAGTGACACTCTTGTCACAGCAG 60  
 Db 5004 TGGGAGAGCTTCACAGGCCCTACCCATATAGTGACACTCTTGTCACAGCAG 5063  
 QY 61 CAGGGAGGACAATTCCCTTACTGGGGGACCGCTAGTAGTGTCAGGATGAGCC 120  
 Db 5064 CAGGGAGGACAATTCCCTTACTGGTGACATGACAGCCAGGTGCGGCAGACT 5123

QY	1321	AAGGCTCTGCCGCGAATTACCGGAGGCCCCATTTCATGCCAACGTGGTAAAGGG 138
Db	6204	AAAGGCTCCACCAAGGATTAACCGGAGACTCTCCACACCGTTCGGCTGGGCTA 6266
QY	1261	AGGGTGTGTTGGAGATGATGCAAGATGTTGGCTGACTTCAGACCTGGCTCAGGCC 132
Db	6264	AGGGTGTGTTGGAGATGATGCAAGATGTTGGCTGACTTCAGACCTGGCTCAGGCC 632
QY	6324	AAGCTCTGCCGCGAACTTACCTTGAGAAGCTTTCCTGTCGACCGGGTAAAGGA 638
QY	1381	GTCCTGGCGGGAGAGGGCATCATGAGAACCCCTGCCATGTGGAGACAGATACCGA 144
Db	6384	GTCCTGGCGGGAGAGGGCATCATGAGAACCCCTGCCATGTGGAGACAGATACCGA 644
QY	1441	CATGTCAAAAAACGGTTCATGAGGATGTTGGCTAGACCTGTCATGTGGAGACAGATACCGA 144
Db	6444	CATGTCAAAAAACGGTTCATGAGGATGTTGGCTAGACCTGTCATGTGGAGACAGATACCGA 650
QY	1501	GGAAACATGCCATAACGATAACCCACGGGCCCCCTGACCCCTTCCCACGCCAAC 156
Db	6504	GGAAACATGCCATAACGATAACCCACGGGCCCCCTGACCCCTTCCCACGCCAAC 656
QY	1561	TATTCAGGGGCTGGGGGGTGGCTGCTGAGGAGTAGTGTGGAGGTTACGGGGTGGG 162
Db	6564	TATTCAGGGGCTGGGGGGTGGCTGCTGAGGAGTAGTGTGGAGGTTACGGGGTGGG 662
QY	1621	GATTTCACATAGCTGTAGGACATGACCACTGACAAGTAAATGCGCTGGCTAGGTTACGGGGTGGG 168
Db	6624	GATTTCACATAGCTGTAGGACATGACCACTGACAACCTAACTGCGCTGGCTAGGTTACGGGGTGGG 668
QY	1681	GCCCCGAACTTCACAGAGTGTGGGGTGGCTCACAGTAGCTGGCTGGGTC 174
Db	6684	GCTCTCTGAAATCTCTGCGAGGTGTGACGGAGTCGGCTGGGTGACAGTGCTCGGCCTGC 674
QY	1741	AAACCTCTCTACGGGAGGGTCACTTCCAGGTGCGGCTAACCAAATACCTGTGGTGG 180
Db	6744	AAACCTCTCTACGGGAGGGTCACTTCCAGGTGCGGCTAACCAAATACCTGTGGTGG 680
QY	1801	TCGAGCTCCATGGAGCCGAGACGGATCTAGCAGTGTCTACTTCATGTCACCGAC 186
Db	6804	TCGAGCTCCATGGAGCCGAGACGGATCTAGCAGTGTCTACTTCATGTCACCGAC 686
QY	1861	CCCTCCCACATCAGCGAGGAGCTAACGGCGCTGGCCAGGGCTCCCCCTCC 192
Db	6864	CCCTCCCACATCAGCGAGGAGCTAACGGCGCTGGCCAGGGCTCCCCCTCC 692
QY	1921	TTGGCAGCTTCAGCTAGCGAAGTGTCTGGCTCTCTCGAGGACATACTTAC 198
Db	6924	TTGGCAGCTTCAGCTAGCGAAGTGTCTGGCTCTCTCGAGGACATACTTAC 698
QY	1981	CAAATGACTTCGGAGACSTGACTTCATCGAGGCCACCTCTGGCGGCTAGAGATG 204
Db	6984	CAAATGACTTCGGAGACSTGACTTCATCGAGGCCACCTCTGGCGGCTAGAGATG 704
QY	2041	GGC 2043
Db	7044	GGC 7046
RESULT	11	
XX	AAD2531	
DE	AAD25321 standard; CDNA, 7987 BP.	
XX	AC	
AC	AAD25321;	
XX	DT	
12-MAR-2002	(first entry)	

to the virus, comprise non-naturally occurring viral sequences.

Db	3577	CTGCACAGGGCCAA GCCCTCTCTGTATAGGCTGGGGCCTTCAAACGGGTTACTAC	3636	Qy	1321	AAGCTCTGGCGGATACGGGAGTCCCTTCTATGCCAACGGGTAAGGG	1380
Qy	241	ACACACCCATAACCAATTCTCATGGCATATCACGGCAGCTGGAGTCGTCACG	300	Db	4717	AAGCTCTGGCGGATTCGGGATTCGGGACTCCCTCTCATGTCAGTGCGGAGA	4776
Db	3637	ACACACCCATAACCAATTCTCATGGCATATCACGGCAGCTGGAGTCGTCACG	3696	Qy	1381	GTCTGGGGAGAGGGCATCATGCAAACCCCTCCATGGAGGAGATCACCGA	1440
Qy	301	AGCACCTGGGCTCTGGGGGCTCTGGGGGCTCTGGGGGCTCTGGGGGCTCTGGGG	360	Db	4777	GTCTGGGGAGAGGGCATCATGCAAACCCCTCCATGGAGGAGATCACCGA	4836
Db	3697	AGCACCTGGGCTCTGGGGGCTCTGGGGGCTCTGGGGGCTCTGGGGGCTCTGGGG	3756	Qy	1441	CATGCAAAACGGGTCATGAGGATGTTGGCTAGAGCTGTAGTACATGTCAT	1500
Qy	361	GCGAGGGGTGTCATGTTGGGGTAGGATCATCTGTCCGGGCCCCGCTATGTTCCGAC	420	Db	4837	CATGCAAAACGGGTCATGAGGATGTTGGCTAGAGCTGTAGTACATGTCAT	4896
Db	3757	GCGAGGGGTGTCATGTTGGGGAGGATCATCTGTCCGGGAAGGCCGCCCATACTCCGAC	3816	Qy	1501	CGAACATTCCTACAAACATACACACGGGCCCCCTGACGCCCTCCAGGCCAAAC	1560
Qy	421	AGGGAAGTCTCTTACCAAGGAGCTGATGAGGAGGAAGAGTCGCGCTGCACTCCCTAC	480	Db	4897	GGAACTTCCCATTAACGGTACACCAACGGGCCCTGCAAGCCCTCCGGCCAAAT	4956
Db	3817	AGGGAAGTCTCTTACCAAGGAGCTGATGAGGAGGAAGTACAGCTGCGCTGCACTCCCTAC	3876	Qy	1561	TATTCAGGGGCTGTCGGCGGTGACTGAGGAGTCTGGAGGTACGGGGTGGGG	1620
Qy	481	ATCGAGGAGGAAATGAGCTGGCGAGCACTGAGGAGTGGCTGAGGAGTGGCTGAG	540	Db	4957	TATTCAGGGGCTGTCGGCGGTGACTGAGGAGTCTGGAGGTACGGGGTGGGG	5016
Db	3877	ATCGAGGAGGAAATGAGCTGGCGAGCACTGAGGAGTGGCTGAGGAGTGGCTGAG	3936	Qy	1621	GATTCACTTACGAGCATGACCACTGACAACTGAAATGCCGGTCCAGTTCCA	1680
Qy	541	ACAGGACCAACCAAGGGGAGGGCGCTGCTCCGGTGGGGTCCAGTGGGAGCCCT	600	Db	5017	GATTCACTTACGAGCATGACCACTGACAACTGAAATGCCGGTCCAGTTCCA	5076
Db	3937	ACAGGACCAACCAAGGAAAGGGAGCTGCTCCGGTGGGGTCCAGTGGGAGCCCT	3996	Qy	1681	GCCCCGAACTTCTCACAGAGTGGATGAGGGCTGAGGGCTGAGGGTACGGGG	1740
Qy	601	GAGACCTTCTGGCGAACACAGATGGGACTCTCATCGCGGATACTGACTTAGAGG	660	Db	5077	GCCCCGAACTTCTCACAGAGTGGATGAGGGCTGAGGGCTGAGGGTACGGGG	5136
Db	3997	GAAGCTTCTGGCGAACACAGATGGGACTCTCATCGCGGATACTGACTTAGAGG	4056	Qy	1741	AAACTCTCTACGGGGAGGTCACTTCAGGGGGTCAACCAACACTGGTGG	1800
Qy	661	TATGCCACTCTGGGAAACACGGGACTCATCTGAGGGCTGACGGCTGTCACG	720	Db	5137	AAACCCCTCTACGGGGAGGTCACTTCAGGGGGTCAACCAACACTGGTGG	5196
Db	4057	TATGCCACTCTGGGAAACACGGGACTCATCTGAGGGCTGACGGCTGTCACG	4116	Qy	1801	TCGCACTCCATGGGAGGCCAACGGGATGAGGTGAGGAGTGGCTCAACT	1860
Qy	721	ACTAGCCGCTTACCAACCAATCTACCTCTGCTAACATCTGGGGGATGGTAGCC	780	Db	5197	TCACGCTCCATGGGAGGCCAACGGGAGTGGCTGACTTCATGCAACCGAC	5256
Db	4117	ACCAGCCGCTTACCAACCAACATACCTCTGGTAAACATCTGGGGGATGGGG	4176	Qy	1861	CCCTCCACATCACAGAGAACGGGAGCTGAGGGCTGAGGGCTGAGGG	1920
Qy	781	GCCCCAACTCGTCCCCAACGGGCTGCTCACTTCTGGTAGGGCTGAGGGCTGAGG	840	Db	5257	CCCTCCACATCACAGAGAACGGGAGCTGAGGGCTGAGGGCTGAGGG	5316
Db	4177	GCCCAACTTCTCCAGGACTGCTGCTTCAGGGCTGAGGGCTGAGGG	4236	Qy	1921	TTGGCAACTTCACTACCACTGGGGAGGTCACTTCAGGGCTTCTGGG	1980
Qy	841	GCTGTTGGGACACATAGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	900	Db	5317	TTGGCAACTTCACTACCACTGGGGAGGTCACTTCAGGGCTTCTGGG	5376
Db	4237	GTTGGTGGGACGATAGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	4296	Qy	1981	CAAATGACTTCCAGAGGCTGACCTTCTGGGGGGGGGGGGGGGGGGGG	2040
Qy	901	GGAGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	960	Db	5377	CGTCACTTCCCGGGAGGTGACTCTATGGGGCAACTCTCTGGGGGGGG	5436
Db	4297	GGGGTGGCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	4356	Qy	2041	GAC	2043
Qy	961	GACCTGGTTAACTTACTCCCTCCATCTCTCTGGGGCTGGGGGGGGGG	1020	Db	5437	GGC	5439
Db	4357	GACTGGTTAACTTACTCCCTCCATCTCTGGGGGGGGGGGGGGGGGGGG	4416	RESUL T 12			
Qy	1021	TGGGCAAGGAGTACTGCGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1080		AAA98968		
Db	4417	TGGGCAAGGAGTACTGCGCTGGGGGGGGGGGGGGGGGGGGGGGGGG	4476		AAA98968	standard; DNA; 7989 BP.	
Qy	1081	CGCTGTAGGAGGCTCGGGGGTAACCATGTTCTCCACGGGACTATGCGAG	1140		AAA98968;		
Db	4477	CGCTGTAGGAGGCTCGGGGGTAACCATGTTCTCCACGGGACTATGCGAG	4536		08-FEB-2001	(first entry)	
Qy	1141	ACGACACGGGAGACGGTGTACTCTGAGCTCTCCACCTACTACGGGACTATGCGAG	1200		XX		
Db	4537	ACGACACGGGAGACGGTGTACTCTGAGCTCTCCACCTACTACGGGACTATGCGAG	4596		XX		
Qy	1201	AAGAGGGCTTCAACGGGACTCTGGGGGGGGGGGGGGGGGGGGGGGGGG	1260		XX		
Db	4597	AAGAGGGCTTCAACGGGACTCTGGGGGGGGGGGGGGGGGGGGGGGGGG	4656		XX		
Qy	1261	ACGGAGCTTCTGGGACTCTGGGACTCTGGGGGGGGGGGGGGGGGGGGGG	1320		XX		
Db	4657	AGAGATGTTGGGATGACGACTCTGGGGGGGGGGGGGGGGGGGGGGGG	4716		XX		



QY	1621	GATTCTCACTTACGGGAGATGACACTGACAACATAATGCCCGAGGTTCCA
QY	5017	GATTCTCACTTACGGGAGATGACACTGACAACATAATGCCCGAGGTTCCA
Db	1681	GCCCCGAATTCTTACAGAAGTGATEGGGTGGTCACAGTAGCTCCGGTAC
QY	5137	ARACCCCTCTACGGGGAGGTACATTCCTGGCTCATCAATACCTGTTGG
Db	5077	GCCCCGAATTCTTACAGAAGTGATEGGGTGGTCACAGTAGCTCCGGTAC
QY	1801	TGGCCTCCATGGAGGCCAACCGATGAGCTAGCTGGCTACTTCATGCTACCGAC
Db	5197	TCACAGTCCCATGGAGGCCAACCGATGAGCTAGCTGGCTACTTCATGCTACCGAC
QY	1861	TGGCCTCCATGGAGGCCAACCGATGAGCTAGCTGGCTACTTCATGCTACCGAC
Db	5257	CCCTCCACATTACGGGAGACGGCTAGCTGGCTACTTCATGCTACCGAC
QY	1921	TGGCAGCTCTCAGGCAGTCTGGCTTCAGGGCTCCACATTAC
Db	5317	TTGGCAGCTCTAGCTGGCTTCAGGGCTCCACATTAC
QY	1981	CAAATGACTTCCAGCTGACTCTATGGGCCACCTCTGGGGCATGAGTG
Db	5377	CCTCATGACTTCCGGAGCTGACTCTATCGAGGCCACCTCTGGGGCATGAGTG
QY	2041	GGC 2043
Db	5437	GGC 5439
RESULT 13		
ID	AAD25322	standard; cDNA; 7989 BP.
XX	AAD25322;	
XX	DT	12-MAR-2002 (first entry)
DE	Hepatitis C virus (HCV) repbBartMan/Avail cDNA.	
XX	KW	Hepatitis C virus; HCV; transfaction; infection; virus neutralisation; gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver ss.
XX	OS	Hepatitis C virus.
XX	FH	Location/Qualifiers
FT	Key CDS	1801...7758
FT	misc_feature	/tag= a /product= "HCVrepbBartMan polyprotein"
FT	misc_feature	/tag= b /note= "Nucleotide creating Avail site"
XX	PN	WO200189364-A2.
XX	PD	29-NOV-2001.
XX	PF	23-MAY-2001; 2001WO-US016822.
PR	XX	23-MAY-2000; 2000US-00576989.
PA	(UNIWI ) UNIV WASHINGTON.	
XX	PT	Rice CM, Blight KJ,
XX	DR	WP1; 2002-066155/09.
DR	XX	P-PSDB; AAE1517.
Hepatitis C virus variants having greater transfection efficiency and		

PT  
PT  
XX  
XX  
PS  
PS  
Claim 44; Page 69-71; 174pp; English.

The invention relates to Hepatitis C virus (HCV) variants which include polynucleotides comprising non-naturally occurring HCV sequence and HCV variants that have a transfection efficiency and ability to survive subpassage greater than HCV that have wild-type polyprotein coding regions. The polynucleotides of the invention are useful for identifying a cell line that is permissive for infection with HCV and detecting replication of HCV in cells of the cell line. They are also useful for testing a compound for anti-viral properties and for inhibiting HCV infection. They are also useful for the generation of defined HCV virus stocks to develop in vitro and in vivo assays for virus neutralisation, attachment, penetration and entry, structure/function studies on HCV proteins and RNA elements and identification of new antiviral targets, a systematic survey of cell culture systems and conditions to identify those that support wild-type and variant HCV RNA replication and particle release, production of adaptive HCV variants capable of more efficient replication in cell culture, production of HCV variants with altered tissue or species tropism, establishment of alternative animal models for inhibitor evaluation including those supporting HCV variant replication, development of cell-free HCV replication assays, production of immunogenic HCV particles for vaccination, engineering of attenuated HCV derivatives as possible vaccine candidates, engineering of attenuated or defective HCV derivatives for expression of heterologous gene products for gene therapy and vaccine applications and for utilisation of the HCV glycoproteins for targeted delivery of therapeutic agents to the liver or other cell types with appropriate receptors. Vaccine comprising these sequences is useful for inducing immunoprotection to HCV in a primate. The present sequence is Hepatitis C Virus (HCV) repBactMan/AvailI CDNA

Db	3877	ATCGAAGCAGGAAATGCAGCTCCGGAAACAACTCAACAGAGGAAATCGGTGCTCAA	3936	Oy	1621	GATTCACTACTGAGGAGCATGACCACTGACAACGTAATGCCGTGCAAGTTCA	1680
Oy	541	ACAGCCACCAAGCAASCGGAGAACCGCGTGTCTCGTGTTGGTCAGTGGAGCCTT	600	Db	5117	GATTCACTACTGAGGAGCATGACCACTGACAACGTAATGCCGTGCAAGTTCA	5076
Db	3937	ACAGCCACCAAGCAACAGGAGGAGCTGCTCCGTTGGACAGGAACTGAGGACCTC	3996	Oy	1681	GCCCCCAATTCCTACAGAAGTGGATGGGTTGGGGTGCGAGGAACTGACGTTG	1740
Oy	601	GAGACCTCTGGGAACACATGTGGAACTCATCGGGATACTGACTTACGGGAAAC	660	Db	5077	GCCCCCAATTCCTACAGAAGTGGATGGGTTGGGGTGCGAGGAACTGACGTTG	5136
Db	3997	GAAGCCATTCTGGGGAGACATATGTGAATTTCATAGCGGGATACTGAGGAAAC	4056	Oy	1741	AAACTCTCTACGGGGAGTCACCTCCAGGTCAGGGGGTCACAGCTCTGTC	1800
Oy	661	TGTGCACTCTGGGAATCCGGGATTCGGGATTCAGTCACTGATGGGTGACAGGTC	720	Db	5137	AAACCTCTACGGGGAGTCACCTCCAGGTCAGGGGGTCACAGCTCTGTC	5196
Db	4057	TGTGCACTCTGGGAACCCGGGATTCAGTCACTGATGGGTGACAGGTCACAGGTC	4116	Oy	1801	TGCACTGCTCCATGGGGAGGAGCTGAGGTCAGTGGCTCACTTCATGTCACCGAC	1860
Oy	721	ACTAGCCGCTTACCAACCAACTACCTCTCTGCTACATGACCTGGGGAGGAGC	780	Db	5197	TCACAGCTCCATGGGGAGGAGCTGAGGTCAGTGGCTCACTTCATGTCACCGAC	5256
Db	4117	ACCAACCGCTTACCAACCAACTACCTCTCTGCTACATGACCTGGGGAGGAGC	4176	Oy	1861	CCCTCCACATCACAGAGGAGGAGCTGAGGTCAGGGGGTCACAGCTACATGTC	1920
Oy	781	CCCAACTCTGGCTCCCAAGCTGCTCTGCTACATGACCTGGGGAGGAGCTG	840	Db	5257	CCCTCCACATCACAGAGGAGGAGCTGAGGTCAGGGGGTCACAGCTACATGTC	5316
Db	4177	GCCTCAACTTGTCTCCAGGGCTGCTCTCTGCTCTGCTTCAGGGGGAGGAGC	4176	Oy	1921	TGGCCGCTCTCACAGGAGGAGCTGAGGTCAGGGGGTCACAGCTACATGTC	1980
Oy	841	GCTGTGCGAGCATAGGCTTGGGAAGGTGCTGTGGACATCTGGGGGATGGAGA	900	Db	5317	TGGCCGCTCTCACAGGAGGAGCTGAGGTCAGGGGGTCACAGCTACATGTC	5376
Db	4237	GCTGTGCGAGCATAGGCTTGGGAAGGTGCTGTGGAGGATTTGGGGAGGATGGAGA	4296	Oy	1981	CAAATCACTTCCAGAGGCTGACCTCATGAGGCCACCTCTGTGGGGCATGATG	2040
Oy	901	GGAGTGCAGGGGGCTGGCTGGCTTAAGGTCATGAGGGGAGATGGGGGATGGAGA	960	Db	5377	CCTCATGACTCCGGAGCTGACCTATCGGGCAACCTCTGGGGCAGGAGATG	5436
Db	4297	GGGTGGCAGGGGGCTGGCTGGCTTAAGGTCATGAGGGGAGATGGGGGATGGAGA	4356	Oy	2041	GTC 2043	
Oy	961	GACCTGGTAACTACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1020	Db	5437	GTC 5439	
Db	4357	GACCTGGTAACTACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4416				
Oy	1021	TGGCAAGGAACTACTGCTCGGACGGTGGCTCAGGGAGGGGGCTGAGGGATGAC	1080				
Db	4417	TGGCAAGGAACTACTGCTCGGACGGTGGCTCAGGGAGGGGGCTGAGGGATGAC	4476				
Oy	1081	CGCTGTAGGTTCTCCCTCCGGGTTAACCTGTTTCCCACGCACTATGTCAGAG	1140				
Db	4477	CGCTGTAGGTTCTCCCTCCGGGTTAACCTGCTCTCCACGGCACTATGTCAGAG	4536				
Oy	1141	AGGAGGCCGCGAGCACTCTGGCTAGATCTCCGACCTTAACCTAACCTAACCTG	1200				
Db	4537	AGGAGGCCGCGAGCACTCTGGCTAGATCTCCGACCTAACCTAACCTAACCTG	4596				
Oy	1201	AGAGGGCTCAACAGGGGATTAACGGGAGCTCTCCAGGCCCTGTCGGGTGCGCTA	1260				
Db	4597	AGAGGGCTCAACAGGGGATTAACGGGAGCTCTCCAGGCCCTGTCGGGTGCGCTA	4656				
Oy	1261	AGGGAGTTCTGGGACTGGATATGCAAGTTGGCTGACTCTCAAGCTGCTCCAGTC	1320				
Db	4657	AGAGGGCTCAACAGGGGATTAACGGGAGCTCTCCAGGCCCTGTCGGGTGCGCTA	4716				
Oy	1321	AGGCTCTGGGGGATTAACGGGAGCTCTCCAGGCCCTGTCGGGTGCGCTAAGGGG	1380				
Db	4717	AGGCTCTGGGGGATTAACGGGAGCTCTCCAGGCCCTGTCGGGTGCGCTAAGGGG	4776				
Oy	1381	GTCTGGGGGAGACGGCATATGAGGACACCTGCTGCTATGGGGACAGATCACGGG	1440				
Db	4777	GTCTGGGGGAGACGGCATATGAGGACACCTGCTGCTATGGGGACAGATCACGGG	4836				
Oy	1441	CATGTCAAACAGGTTCATGGATGTTGGCTTAAGGCTAACCTGGCTTGGCTAAG	1500				
Db	4837	CATGTCAAACAGGTTCATGGATGTTGGCTTAAGGCTAACCTGGCTTGGCTAAG	4896				
Oy	1501	GGAAACATTCCTACAGCTAACACGGGCCCCCTGGAGGCTTGGCTAAGGGGAGG	1560				
Db	4897	GGAAACATTCCTACAGCTAACACGGGCCCCCTGGAGGCTTGGCTAAGGGGAGG	4956				
Oy	1561	TATTCAGGGGCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1620				
Db	4957	TATTCAGGGGCTGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5016				

WPI: 2004-18065/17.

P-PSDB; ADJ57846.

XO (BRIM ) BRISTOL-MYERS SQUIBB CO.

P Gao M, Lemm JA, O'boyle DR, Nower P, Rigat K, Sun J;

PR 12-AUG-2003; 2003WO-US025260.

XX 12-AUG-2002; 2002US-040261P.

XX Use of hepatitis C virus assays or reporter assays, e.g. identifying a

PT compound that inhibits hepatitis C virus RNA replication or identifying a

PT compound that modulates the activity of a gene of interest.

QY	4057	TTCGTCACCTCGCGTGCACCCGGATAGCATCACTGAGGCACTACAGCCTAAC	4116
Db	721	ACTAGCCCCGTCAACCCAAATCTACCCCTCTGCTTAACATCTCGGGATGGTAGCC	780
Db	4117	ACCAGCCCTCACCAACACATACCCCTCTGCTTAACATCTCGGGATGGTAGCC	4176
QY	781	GCCCCAACTCTCCACCCCGAGTGTGCTGAAGGAGCTTGAGGAGTTACCGGGGGCG	840
Db	4177	GCCCCAACTCTCTCCACAGCCTCTCTGCTTGAGGAGCTTGAGGAGTTACCGGGGGCG	4236
QY	841	GCTGTTGSCAGCATAGGCCTTGCGAGGAGCTTGAGAATCTCTGGGGATGGAGGA	900
Db	4237	GCTGTTGSCAGCATAGGCCTTGCGAGGAGCTTGAGAATCTCTGGGGATGGAGGA	4296
QY	901	GGAGTGCGAGGGGGCTCGGGCTTAAGGTCTATGAGCGGGGAATGGCTCCACCGAG	960
Db	4297	GGGTGCGAGGGGGCTCGGGCTTAAGGTCTATGAGCGGGGAATGGCTCCACCGAG	4356
QY	961	GACCTGTTAATRACTACTCCCTGCCATTCCCTCTCTCTGGGCCCTGTGGACTGGAG	1020
Db	4357	GACCTGTTAATRACTACTCCCTGCCATTCCCTCTCTCTGGGCCCTGTGGACTGGAG	4116
QY	1021	TGCGGAGCGATACTCGTCGGACAGTGGGGTCCAGGGAGGGGTGGACTGGAGAAC	1080
Db	4417	TGCGGAGCGATACTCGTCGGACAGTGGGGTCCAGGGAGGGGTGGACTGGAGAAC	4476
QY	1081	CGGCTGATAGCGTTCGGCTCGCGGGGGTAACCATGTTTCCCACGGACTATGTCAGAG	1140
Db	4477	CGGCTGATAGCGTTCGGCTCGCGGGGGTAACCATGTTTCCCACGGACTATGTCAGAG	4536
QY	1141	AGCCAGCGCAGACAGTGTCACTCGAGCTCTCGGACCTTAATCACCCAACTGTG	1200
Db	4537	AGCCAGCGCAGACAGTGTCACTCGAGCTCTCGGACCTTAATCACCCAACTGTG	4596
QY	1201	AAGAGGCTCACCAGTGGATTACCGGACTGTGTCACATCGAGCTCTCGGACCTTAATCACCGTCA	1260
Db	4597	AAGAGGCTCACCAGTGGATTACCGGACTGTGTCACATCGAGCTCTCGGACCTTAATCACCGTCA	4656
QY	1261	AGGGATGTTGGACTGTGGATATGCAAGTTGGGTACTTCAGCTGGCTCAGTCC	1320
Db	4657	AGGAGTGTGTTGGATTATGAGGATGAGGGTGTACTGATGTTCAAGGACTGGCTCCAGTCC	4716
QY	1321	AGCTCTGGGATAACGGGACTCCCTTTCTCTGCAAGCTGGTGTACAGGGG	1380
Db	4717	AGCTCTGGGATAACGGGACTCCCTTTCTCTGCAAGCTGGTGTACAGGGG	4776
QY	1381	GTCTCGGGGGAGACGGCATCATGAGACACCTCTCATGTGAGCACAGATCGGGA	1440
Db	4777	GTCTCGGGGGAGACGGCATCATGAGACACCTCTCATGTGAGCACAGATCGGGA	4836
QY	1441	CATGTCAAAACGGTTCCATGAGGATCTGGGGCTTAAGACCTGTTGAGTAACTGTCGAT	1500
Db	4837	CATGTCAAAACGGTTCCATGAGGATCTGGGGCTTAAGACCTGTTGAGTAACTGTCGAT	4896
QY	1501	GGAACATCCCATCACGGATACACACGGGCTCTGGCACGCCCTCCAGGSCAAC	1560
Db	4897	GGAACATCCCATCACGGATACACACGGGCTCTGGCACGCCCTCCAGGSCAAC	4956
QY	1561	TATTCAGGGCGCTGGGGGTGGTGTGAGGAGGTACGGTGGAGGTTACGGGGGGGG	1620
Db	4957	TATTCAGGGCGCTGGGGGTGGTGTGAGGAGGTACGGTGGAGGTTACGGGGGGGG	5016
QY	1621	GATTCACACTACGTCAGGAGCATGACACTGACAAGTAATGCCGCAAGGTCCA	1680
Db	5017	GATTCACACTACGTCAGGAGCATGACACTGACAAGTAATGCCGCAAGGTCCA	5076
QY	1681	GCCCCGAATCTCACAGAGTGTGGGGTGGGGCTGCGCTGAAGGTACGGTCCGGGTG	1740
Db	5077	GCCCCGAATCTCACAGAGTGTGGGGTGGGGCTGCGCTGAAGGTACGGTCCGGGTG	5136
QY	1741	AAACCTCTACGGGGAGGTACCATTCAGGAGGGCTCAACAAACTCTGGTGG	1800
Db	5137	AAACCTCTACGGGGAGGTACCATTCAGGAGGGCTCAACAAACTCTGGTGG	5196

QY 1801 TCCGAGCTTCCATGCGAGCCCCGAAACGGATGTA3CAGTGCTCACTTCATGTCACCGAC 1860  
Db 5197 TCACAGTCATCGAGGCCGAAACGGACTGCTGACTTCATGTCACCGAC 5256  
QY 1861 CCCCCTCATCACAGAGAGGGCTRAAGCCAGGTAGCCAGGCTGCTACTTCATGTCACCGAC 1920  
Db 5257 CCCCTCCAAATTAGGGGAAGGGTAAGCGTAGGCTGGCCAGGGATCCTCCCTCC 5316  
QY 1921 TTGGCCAGCTCTCAGTAGGGTGTGCTGCAGGACTCTTGAGGGACTACATTAAC 1980  
Db 5317 TTGGCCAGCTCTCAGTAGGGTGTGCTGCAGGACTCTTGAGGGACTACATTAAC 5376  
QY 1981 CAAATGACTTCCAGAGGCTGACTCATGGAGCCAACTCTTGAGGGATGAGATG 2040  
Db 5377 CCTCATGACTCCCGAGCTGACTCTGAGCCAACTCTTGAGGGAGATG 5436  
QY 2041 GSC 2043  
Db 5437 GGC 5439

RESULT 15  
ID AAL4726  
AAL4726 standard; DNA; 7992 BP.  
XX  
AC AAL4726;  
XX  
DT 30-AUG-2002 (first entry)  
DE Hepatitis C virus sub-genomic replicon clone 1377-NS3-3' UTR.  
XX  
KW Hepatitis; HCV; core-neo; NS3 proteinase; vaccine; diagnosis;  
viricide; hepatotropic; gene therapy; anti-viral; gene; ds.  
XX  
OS Hepatitis C virus.

Key Location/Qualifiers  
FH CDS 342. .1181  
FT /\*tag= a  
FT /product= "core-neo fusion protein"  
FT .7758  
FT /\*tag= b  
FT /product= "NS3 proteinase/helicase"  
XX  
WO200238793-A2.  
PD 16-MAY-2002.  
XX  
PP 02-NOV-2001; 2001WO-US046350.  
XX  
PR 07-NOV-2000; 2000US-0245866P.  
XX  
PA (ANAD-) ANALYS PHARM INC.  
XX  
PI Bichko V;  
DR WPI; 2002-490082/52.  
PT P-PSDB; A018000, A018001.  
XX  
The present invention provides protein and coding sequences from  
Hepatitis C virus (HCV), comprising all or part of the HCV genome and  
able to replicate efficiently when transfected into a susceptible cell  
line without reducing the growth rate of the cell line by more than 10  
fold. The sequences are useful for screening for anti-HCV therapeutics,  
for detecting antibodies to HCV in a biological sample such as blood,  
serum, plasma, blood cells, lymphocytes, or liver cells from a subject,

CC for deriving authentic HCV components such as replication-complement non-  
infectious, replication-defective infection-component, and replication-  
defective non-infectious HCV, in gene therapy or gene vaccination  
targeted to hepatic tissue for treating an animal infected or susceptible  
to HCV infection and for studying HCV infection and propagation. The  
present sequence is a clone of a fragment of the HCV genome which encodes  
the core-neo and NS3 proteinease/helicase proteins

XX Sequence 7992 BP; 1648 A; 2369 C; 2243 G; 1732 T; 0 U; 0 Other;  
Query Match Best Local Similarity 87.5%; Score 1787; DB 6; Length 7992;  
Matches 1883; Conservative 92.2%; Pred. No. 0;  
MisMatches 160; Index 0; Gaps 0;

Db 3397 TGGAGAGCGCTTTACAGGCTTACCCACATAGACGCCATTCTGTCAGGACTAAG 3456  
QY 61 CAGGGAGAGACAATCCCTTACCTGGTGGTACCCAGGTACTCTGTGCTGGCTAGGG 120  
Db 3457 CAGGAGAGACAATCCCTTACCTGGTGGTACCCACATAGACGCCATTCTGTCAGGACTAAG 3516  
QY 121 CAGGCCAACCTTCAATGGATCAAATGGGAAGTGTGCTCATGGCTAAAGCTACT 180  
Db 3517 CAGGCTAACCTTCAATGGGAACCAATGGGAAGTGTGCTCATGGCTAAAGCTACT 3576  
QY 181 CTGGAGGGCGCTTCACACCCCTGTATAGGGTGGAGGCCAAAGAGAGTCACCC 240  
Db 3577 CTGACACGGCCAAACGCCCTGTGTAAAGCTGGAGGCCAAAGAGAGTCACCC 3636  
QY 241 ACACACCCATAACAAATCATCGATCGATCATCGACCTGGAGCTGGCTGCG 300  
Db 3637 ACACACCCATAACAAATCATCGATCGACCTGGAGCTGGCTGCG 3696  
QY 301 AGGACCTGGTCTGGGGGGCTTGAGCTGGAGCTGGCTGCGTGTGAGCTGG 360  
Db 3697 AGCACCGAGGCTGGAGGGAGCTGAGCTGGAGCTGGCTGCGTGTGAGCTGG 3756  
QY 361 GGCAGCTGGTCTGGGGGGAGCTGAGCTGGCTGCGTGTGAGCTGGCTGCG 420  
Db 3757 GGCAGCTGGTCTGGGGGGAGCTGAGCTGGCTGCGTGTGAGCTGGCTGCG 3816  
QY 421 AGGGAGTCTCTACGGGAGTCGATAGTGGAGAGTCGGCTGGCTGGAGCTGG 480  
Db 3817 AGGGAGTCTCTACGGGAGTCGATAGTGGAGAGTCGGCTGGCTGGAGCTGG 3876  
QY 481 ATCGAGGAGGGATCGAGCTCCGAGCAGTCAGCAAAGGGCTGGGTGCTCAG 540  
Db 3977 ATCGAGGAGGGATCGAGCTCCGAGCAGTCAGCAAAGGGCTGGGTGCTCAG 3936  
QY 541 ACAGCCCAAGCAAGGGAGGCTGCTCCGAGTCAGTGGAGCTGGCTGGAGCCT 600  
Db 3937 ACAGCCCAAGCAAGGGAGGCTGCTCCGAGTCAGTGGAGCTGGCTGGAGCCT 3996  
QY 601 GAGACCTCTGGGAAACACATGGGAGACTCATGGCGATCAGTACTTACAGG 660  
Db 3997 GAGGCTCTGGGAGCATATGTTGAAATTACATGCGGGATACATATTAGGGC 4056  
QY 661 TTGTCACCTCTGGGAAACACATGGGAGACTCATGGCGATCAGTACTTACAGG 720  
Db 4057 TTGGCTCTGGGAGCATATGTTGAAATTACATGCGGGATACATATTAGGGC 4116  
QY 721 ACTAGCCGCTTACACCAACTTACCTCTCTGGTACATTCCTGGGGATGGTAGGC 780  
Db 4117 ACCAGCCGCTTACACCAACTTACCTCTGGTACATTCCTGGGGATGGGG 4176  
QY 781 GCGCAACTCTGGGAGCATATGTTGAAATTACATGCGGGATACATATTAGGGC 840  
Db 4177 GCGCAACTCTGGGAGCATATGTTGAAATTACATGCGGGATGGGG 4236  
QY 841 GCTGTTGGCAGATAGGGCTGGGAAGGGTGTGAGACATCTGGGGGTATGGAGCA 900  
Db 4237 GCTGTTGGCAGATAGGGCTGGGAAGGGTGTGAGACATCTGGGGGTATGGAGCA 4296

Qy 901 CGAGTGGCAGGCGCCCTGTGGCTTAAGTCATGAGCGGAATGCCCTCACCGAG 960  
Db 4297 GGGTGGCAGCGCGCTGTGGCTTAAGTCATGAGCGGAATGCCCTCACCGAG 4356  
Qy 961 GACCTGGTTAACCTACTTACCTCCCTCCATCTCTCTCTGGCTGCAGGGAG 1020  
Db 4357 GACCTGGTTAACCTACTTACCTCCCTCCATCTCTCTCTGGCTGCAGGGAG 4416  
Qy 1021 TCGCGACGGATACTGCTCGGACGCTGGGTCAAGGGAGGGCTGAGTGAC 1080  
Db 4417 TCGCGACGGATACTGCTCGGACGCTGGGTCAAGGGAGGGCTGAGTGAC 4476  
Qy 1081 CGCGTATAGCGTTGCTCGGGTRACCATGTTCCCACGGCATATGTGGAG 1140  
Db 4477 CGCGTATAGCGTTGCTCGGGTRACCATGCTCCACGGCATATGTGGAG 4536  
Qy 1141 AGCGACCCGCGACGACCGTGTACTCAGATCTCTCCACGCTGCGGGTGTG 1200  
Db 4537 AGCGACCCGCGACGACCGTGTACTCAGATCTCTCCACGCTGCGGGTGTG 4596  
Qy 1201 AAGAGSCTCCACAGTGAATACGAGGACTCTCCACGCTGAGTCTACCA 1260  
Db 4597 AAGAGSCTCCACAGTGAATACGAGGACTCTCCACGCTGAGTCTACCA 4656  
Qy 1261 AGGGATTTGGACTGGATTCACAGTTGGCTGACTCAAGACCTGGTCCAGTC 1320  
Db 4657 AGAGATGTTGGATGGATAGTCACCGGTGACTGATTCAAGACTGSGTCCAGTC 4716  
Qy 1321 AAAGCTCTGCGCGGATACCGGGAGTCCCTTTTCATGCCAACCTGGTACAGGG 1380  
Db 4717 AAGCTCTGCGCGGATACCGGGAGTCCCTTTTCATGCCAACCTGGTACAGGG 4776  
Qy 1381 GCTGGGGGGAGACGGCATCAGCAGAACCTGCTATGGAGCACAGATCACCGA 1440  
Db 4777 GCTGGGGGGAGACGGCATTCAGCAGAACCTGCTATGGAGCACAGATCACCGA 4836  
Qy 1441 CATGTCAAAGGGTTCATGAGGATGTTGGCTTAAGACCTGTTGACTAAATGTCAT 1500  
Db 4837 CATGTCAAAGGGTTCATGAGGATGTTGGCTTAAGACCTGTTGACTAAATGTCAT 4896  
Qy 1501 GGACATCCCCATCAAGGATAACCAACGGGCCCCCTGACGCCCTCCGGCCAAAC 1560  
Db 4897 GGACATCCCCATCAAGGATAACCAACGGGCCCCCTGACGCCCTCCGGCCAAAT 4956  
Qy 1561 TATTCAGGGCCTGGGGGGCTGCTGGAGGATAGTGAGGTTACCGGGTGG 1620  
Db 4957 TATTCAGGGCCTGGGGGGCTGCTGGAGGATAGTGAGGTTACCGGGTGG 5016  
Qy 1621 GATTTCCACTCTGAGGATGACCACTGACAACGTTAAATGCCGTGCAAGGTRCA 1680  
Db 5017 GATTTCCACTCTGAGGATGACCACTGACAACGTTAAATGCCGTGCAAGGTRCA 5076  
Qy 1681 GCCCCGGAATCTTCAGAGTGGATGGGTCGGCTGCAAGGGTAGCTCGGGCTG 1740  
Db 5077 GCCCCGGAATCTTCAGAGTGGATGGGTCGGCTGCAAGGGTAGCTCGGGCTG 5136  
Qy 1741 AACCTCTCCATCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 1800  
Db 5137 AACCCCTCTACCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 5196  
Qy 1801 TGGCAGCTCCATCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 1860  
Db 5197 TCAAGCTCCATCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 5256  
Qy 1861 CCTCCCATCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 1920  
Db 5257 CCTCCCATCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 5316  
Qy 1921 TGGCAGCTCTGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 1980  
Db 5317 TGGCAGCTATCGGGAGGGTACATTCAGGGTCAACCAATCTGGTGG 5376

Qy 1981 CAAATGACTTCCACAGCGTGACCTGAGGGCAACCTGCGGCGATGAGTG 2040  
Db 5377 CGTCATGACTCCCGGAGCTGACTCATCGAGGCCACCTCTGAGGAGATG 5436  
Qy 2041 GGC 2043  
Db 5437 GGC 5439  
Search completed: February 25, 2005, 03:32:08  
Job time : 964 secs

Copyright (c) 1993 - 2005 Compugen Ltd.

Om nucleic - nucleic search, using sw model

Run on: February 25, 2005, 03:10:04 ; Search time 366 Seconds

Title: US-09-664-363-20

Perfect score: 2043

Sequence: 1 TGGAGGCGGTCTCACAGG.....TGTGGCCGCATGAGATGGC 2043

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cggn2\_6/ptodata/1/ina/5A COMB. seq: \*  
2: /cggn2\_6/ptodata/1/ina/5B COMB. seq: \*  
3: /cggn2\_6/ptodata/1/ina/6A COMB. seq: \*  
4: /cggn2\_6/ptodata/1/ina/6B COMB. seq: \*  
5: /cggn2\_6/ptodata/1/ina/PCTUS COMB. seq: \*  
6: /cggn2\_6/ptodata/1/ina/backfile1.seq: \*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

**SUMMARIES**

Result No.	Score	Query Match Length	DB ID	Description
1	2043	100.0	2043	3 US-08-191-160-20 Sequence 20, Appl
2	2043	100.0	3750	3 US-08-191-160-22 Sequence 22, Appl
3	1796.6	87.9	7863	1 US-08-324-977-35 Sequence 35, Appl
4	1796.6	87.9	7863	2 US-08-384-636-35 Sequence 35, Appl
5	1796.6	87.9	7863	3 US-08-304-636A-35 Sequence 35, Appl
6	1796.6	87.9	7863	4 US-09-315-830-35 Sequence 31, Appl
7	1796.6	87.9	7917	1 US-08-324-977-31 Sequence 31, Appl
8	1796.6	87.9	7917	2 US-08-384-636-31 Sequence 31, Appl
9	1796.6	87.9	7917	2 US-08-304-636A-31 Sequence 31, Appl
10	1796.6	87.9	7917	3 US-09-315-830-31 Sequence 31, Appl
11	1796.6	87.9	9030	2 US-08-324-977-13 Sequence 13, Appl
12	1796.6	87.9	9030	2 US-08-384-636-13 Sequence 13, Appl
13	1796.6	87.9	9030	3 US-08-304-636A-13 Sequence 13, Appl
14	1796.6	87.9	9030	3 US-09-315-830-13 Sequence 13, Appl
15	1796.6	87.9	9416	1 US-08-324-977-1 Sequence 1, Appl
16	1796.6	87.9	9416	2 US-08-384-636A-1 Sequence 1, Appl
17	1796.6	87.9	9416	3 US-09-315-830-1 Sequence 1, Appl
18	1796.6	87.9	9416	3 US-08-323-855A-27 Sequence 13, Appl
19	1796.6	87.9	9416	3 US-08-323-855A-27 Sequence 10, Appl
20	1796.6	87.9	7989	4 US-09-339-601-10 Sequence 7, Appl
21	1787	87.5	8001	4 US-09-539-601-7 Sequence 4, Appl
22	1787	87.5	8637	4 US-09-539-601-4 Sequence 4, Appl
23	1787	87.5	8638	4 US-10-029-907-24 Sequence 24, Appl
24	1787	87.5	8639	4 US-10-029-907-1 Sequence 1, Appl
25	1787	87.5	8649	4 US-09-539-601-13 Sequence 13, Appl
26	1787	87.5	1076	4 US-09-539-601-2 Sequence 1, Appl
27	1786.6	87.4	8642	4 US-10-029-907-2 Sequence 2, Appl

**ALIGNMENTS**

RESULT 1  
US-08-191-160-20  
; Sequence 20, Application US/08191160  
; Patent No. 6210675  
; GENERAL INFORMATION:  
; APPLICANT: Highfield, Peter Edmund  
; APPLICANT: Rodger, Brian Colin  
; APPLICANT: Tedder, Richard Seton  
; APPLICANT: Barbara, John Anthony James  
; TITLE OF INVENTION: Viral Agent  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSE: Rothwell, Figg, Ernst & Kurz  
; STREET: 1700 K Street  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20006  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy diskette, 5.25 inch, 360 Kb storage  
; COMPUTER: IBM AT compatible  
; OPERATING SYSTEM: MS-DOS V3.2  
; SOFTWARE: Wordperfect 5.0 (DOS text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/191.160  
; FILING DATE:  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/628, 516  
; FILING DATE: 17 DEC 1990  
; APPLICATION NUMBER: UK 89 28 562.1  
; FILING DATE: 18 DEC 1989  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 414.0  
; FILING DATE: 27 FEB 1990  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: UK 90 04 814.1  
; FILING DATE: 03 MAR 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: E. Anthony Figg  
; REGISTRATION NUMBER: 27, 195  
; REFERENCE/DOCKET NUMBER: 1645-103A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 833-5740  
; TELEFAX: (202) 833-5744  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2043 base pairs  
; TYPE: nucleotide with corresponding protein  
; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: cDNA to genomic RNA  
ORIGINAL SOURCE:  
ORGANISM: human; serum infectious for PT-NANBH  
IMMEDIATE SOURCE: PT-NANBH  
LIBRARY: clone 156/92

FEATURE: from 1 to 2043 bp portion of the PT-NANBH  
LOCATION: polyprotein  
OTHER INFORMATION: probably encodes viral non-structural  
; OTHER INFORMATION: proteins

US-08-191-160-20

Query Match	100 %;	Score 2043;	DB 3;	Length 2043;
Best Local Similarity	100 %;	Pred. No. 0;	Mismatches 0;	Gaps 0;
Matches 2043;	Conservative 0;	Indels 0;		

Qy 1 TGGAGGGGCGTTCACAGGCTCACCCACCTGGATGCCACTCTGTCCAAACAAAG 60  
Db 1 TGGAGGGGCGTTCACAGGCTCACCCACCTGGATGCCACTCTGTCCAAACAAAG 60  
Qy 61 CAGGAGGAGAACACTTCCCACCTGGTGGCTACAGGACTACTGTGGCTAGGGCC 120  
Db 61 CAGGAGGAGAACACTTCCCACCTGGTGGCTACAGGACTACTGTGGCTAGGGCC 120  
Qy 121 CAGGCCAACCTCCATCATGGATAATGSGAAGCTGTCATAGGGTAAGCTACT 180  
Db 121 CAGGCCAACCTCCATCATGGATAATGSGAAGCTGTCATAGGGTAAGCTACT 180  
Qy 181 CTGGCGGGGCCAACACCTTGCTGTTAGGTGGGACCCGCTCAAAACGGGTACCC 240  
Db 181 CTGGCGGGGCCAACACCTTGCTGTTAGGTGGGACCCGCTCAAAACGGGTACCC 240  
Qy 241 ACACACCCATACCAATTATCATGGCATCATGTCAGCGACCTGGGAGTCACCTC 300  
Db 241 ACACACCCATACCAATTATCATGGCATCATGTCAGCGACCTGGGAGTCACCTC 300  
Qy 301 AGCACCTGGGACTGTTGGGGGCTTCAGCTTGCTGCGTATTGTCAGAAC 360  
Db 301 AGCACCTGGGACTGTTGGGGGCTTCAGCTTGCTGCGTATTGTCAGAAC 360  
Qy 361 GCGAGCTGGTATGGGGAGCATCTGTCGGGCGCCGCTATGTCGGAC 420  
Db 361 GCGAGCTGGTATGGGGAGCATCTGTCGGGCGCCGCTATGTCGGAC 420  
Qy 421 AGGAAACTCTCTTACCAAGGAGTCGATGAGTGGGAAGAGTCGGCGCACCTCCCTAC 480  
Db 421 AGGAAACTCTCTTACCAAGGAGTCGATGAGTGGGAAGAGTCGGCGCACCTCCCTAC 480  
Qy 481 ATGCAGGAGGATGAGTCGGCGAGCAAGAAAGGCTCGGCTTGCGAC 540  
Db 481 ATGCAGGAGGATGAGTCGGCGAGCAAGAAAGGCTCGGCTTGCGAC 540  
Qy 541 AGCGCCACAGGAGGGGCGCTGTCGGTGGAGTCCAGTGGACGCTT 600  
Db 541 AGCGCCACAGGAGGGGCGCTGTCGGTGGAGTCCAGTGGACGCTT 600  
Qy 601 GAGACCTCTGGCGAACACAGTGGACTCATCGCGGATACAGTAGCTAGGAGC 660  
Db 601 GAGACCTCTGGCGAACACAGTGGACTCATCGCGGATACAGTAGCTAGGAGC 660  
Qy 661 TGGCACTCTGGCGAACACAGTGGACTCATCGCGGATACAGTAGCTAGGAGC 720  
Db 661 TGGCACTCTGGCGAACACAGTGGACTCATCGCGGATACAGTAGCTAGGAGC 720  
Qy 721 ACTAGCGCTTACACCCATCTACCTCTGGTACATCTGGGGATGGGAGCC 780  
Db 721 ACTAGCGCTTACACCCATCTACCTCTGGTACATCTGGGGATGGGAGCC 780  
Qy 781 GCGCAACTCGCTCCCCCAGTCGCTCAAGCTTCTGGCTAGGGCGCATCTGGCG 840  
Db 781 GCGCAACTCGCTCCCCCAGTCGCTCAAGCTTCTGGCTAGGGCGCATCTGGCG 840  
Qy 841 GCTGTTGCGACATAGGCTTGGGAGGAGCTCTGGGGCTATGAGCA 900  
Db 841 GCTGTTGCGACATAGGCTTGGGAGGAGCTCTGGGGCTATGAGCA 900  
Qy 901 GGACTGAGCGGCCTCGGGCTTAAAGTCATGGGGAAATGCCCTCACCGAG 960  
Db 901 GGAGTGGAGCGGCCTCGGGCTTAAAGTCATGGGGAAATGCCCTCACCGAG 960  
Qy 961 GACCTGGTAACTACTCTCTGGCCATCTCTCTGGTGGCTGGGGTGTG 1020  
Db 961 GACCTGGTAACTACTCTCTGGCCATCTCTCTGGTGGCTGGGGTGTG 1020  
Qy 1021 TGGCGCGGAGTACCTGGCGCACCGGGGGCTGGAGTGGATGAC 1080  
Db 1021 TGGCGCGGAGTACCTGGCGCACCGGGGGCTGGAGTGGATGAC 1080  
Qy 1081 CGGCTGATACCGTTCGCTCGGGGGTAACCATGTTCCCAAGGCACTATGTCAGAG 1140  
Db 1081 CGGCTGATACCGTTCGCTCGGGGGTAACCATGTTCCCAAGGCACTATGTCAGAG 1140  
Qy 1141 AGCGACGGCCAGCACGCTGTCACTCAGATCTCTCGGACTTACTATACCCAACTGTG 1200  
Db 1141 AGCGACGGCCAGCACGCTGTCACTCAGATCTCTCGGACTTACTATACCCAACTGTG 1200  
Qy 1201 AAGAGCTCCACCGAGGATTAACAGGAGCTGCTCAAGGCTGTGCTGAGCTG 1260  
Db 1201 AAGAGCTCCACCGAGGATTAACAGGAGCTGCTCAAGGCTGTGCTGAGCTG 1260  
Qy 1261 AGGGATGTTGGACTGATACTGCACTGAGTTTGCTGTTGACTCTCAAGAC 1320  
Db 1261 AGGGATGTTGGACTGATACTGCACTGAGTTTGCTGTTGACTCTCAAGAC 1320  
Qy 1321 AAGCTCTGGGGATTCCGGAGTCCTCTTCTATGCCAACGTCAGTGGTACAGGG 1380  
Db 1321 AAGCTCTGGGGATTCCGGAGTCCTCTTCTATGCCAACGTCAGTGGTACAGGG 1380  
Qy 1381 GTCTGGGGAGACGGCATGAGGACACCTCTCATGGAGACAGATACCGGA 1440  
Db 1381 GTCTGGGGAGACGGCATGAGGACACCTCTCATGGAGACAGATACCGGA 1440  
Qy 1441 CATGTCAAACAGGTTCACTGAGGATGTTGGCTCTAGACCTGTTGAC 1500  
Db 1441 CATGTCAAACAGGTTCACTGAGGATGTTGGCTCTAGACCTGTTGAC 1500  
Qy 1501 GGAACATCTCCATCAAAGGATACCAACGGCCCTCAGGCCCTCCAGGCSAAC 1560  
Db 1501 GGAACATCTCCATCAAAGGATACCAACGGCCCTCAGGCCCTCCAGGCSAAC 1560  
Qy 1561 TAATCGAGGCTGCGGGTGGAGGATGACCTGGGTTAGAGGTACCTGGG 1620  
Db 1561 TAATCGAGGCTGCGGGTGGAGGATGACCTGGGTTAGAGGTACCTGGG 1620  
Qy 1621 GATTCAGCACTAGTGAGGAGCATGACCACTGAGCAACGTAATGCGCTCCAGGTTCA 1680  
Db 1621 GATTCAGCACTAGTGAGGAGCATGACCACTGAGCAACGTAATGCGCTCCAGGTTCA 1680  
Qy 1680 GCGCCGAATCTGGAGGAGCTGGGGTGGAGGATGACCTGGGTTAGAGGTACCTGGG 1740  
Db 1680 GCGCCGAATCTGGAGGAGCTGGGGTGGAGGATGACCTGGGTTAGAGGTACCTGGG 1740  
Qy 1741 AACCTCTCTGGAGGAGGAGCATGACCACTGAGCAACGTAATGCGCTCCAGGTTCA 1800  
Db 1741 AACCTCTCTGGAGGAGGAGCATGACCACTGAGCAACGTAATGCGCTCCAGGTTCA 1800  
Qy 1801 TCGCACTCCATGGGAGGAGGAGCATGACCACTGAGCAACGTAATGCGCTCCAGGAC 1860  
Db 1801 TCGCACTCCATGGGAGGAGGAGCATGACCACTGAGCAACGTAATGCGCTCCAGGAC 1860  
Qy 1861 CCCTCCACATCACGGAGGAGCTAACGGAGGCTGGGGGCTGGGGGCTCC 1920  
Db 1861 CCCTCCACATCACGGAGGAGCTAACGGAGGCTGGGGGCTGGGGGCTCC 1920  
Qy 1920 TTGGCACTCTGGCTAGGAGGAGGAGCATGACCACTGAGCAACGTAATGCGCTCC 1980

Query Match 100.0%; Score 2043; DB 3; Length 3750;  
 Best Local Similarity 100.0%; Pred. No. 0; Matches 2043; Conservative 0; Mismatches 0; Index 0; Gaps 0;

Matches 2043; Conservative 0; Mismatches 0; Index 0; Gaps 0;

Db 1921 TTGGCAGCTTCAGTGTAGCGCTTCGCTTGAGGCACATATTAC 1980  
 Qy 1981 CAAATGACTTCCAGGGCTGACCTCATGGGCCACCTCTGGGCGCATGAGATG 2040  
 Db 1981 CAAATGACTTCCAGGGCTGACCTCATGGGCCACCTCTGGGCGCATGAGATG 2040  
 Qy 2041 GGC 2043  
 Db 2041 GGC 2043

RESULT 2  
 US-08-191-160-22  
 Sequence 22, Application US/08191160  
 Patent No. 6310675  
 GENERAL INFORMATION:  
 APPLICANT: Highfield, Peter Edmund  
 APPLICANT: Rodgers, Brian Colin  
 APPLICANT: Tedder, Richard Seaton  
 APPLICANT: Barbara, John Anthony James  
 TITLE OF INVENTION: Viral Agent  
 NUMBER OF SEQUENCES: 25  
 CORRESPONDENCE ADDRESS:  
 ADDRESSE: Rothwell, Figg, Ernst & Kurz  
 STREET: 1700 K Street  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.

ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Flopy diskette, 5.25 inch, 360 KB storage  
 COMPUTER: IBM AT compatible  
 OPERATING SYSTEM: MS-DOS V3.2  
 SOFTWARE: Wordperfect 5.0 (DOS text)

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/191,160  
 FILING DATE:  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: 07/628,516  
 FILING DATE: 17 DEC 1990  
 APPLICATION NUMBER: UK 89 28 562.1  
 FILING DATE: 18 DEC 1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: UK 90 04 414.0  
 FILING DATE: 27 FEB 1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: UK 90 04 814.1  
 FILING DATE: 03 MAR 1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: E. Anthony Figg  
 REGISTRATION NUMBER: 27,195  
 REFERENCE NUMBER: 1645-103A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 833-5740  
 TELEFAX: (202) 833-5744  
 INFORMATION FOR SEQ ID NO: 22:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3750 base pairs  
 TYPE: nucleotide with corresponding protein  
 STRANDEDNESS: single  
 MOLECULE TYPE: cDNA to genomic RNA  
 ORIGINAL SOURCE:  
 ORGANISM: human; serum infections for PT-NANBH  
 IMMEDIATE SOURCE:  
 LIBRARY: cDNA clones from 3' end of the genome  
 FEATURE:  
 LOCATION: from 1 to 3750 bp portion of the PT-NANBH  
 LOCATION: polyprotein  
 OTHER INFORMATION: viral non-structural proteins  
 US-08-191-160-22

Db 121 CAGGCCAACCTCATGGGATCAATGGGATCATGGGAAGTCTCATGGCTAACGCT 180  
 Qy 121 CAGGCCAACCTCATGGGATCAATGGGATCATGGGAAGTCTCATGGCTAACGCT 180  
 Db 121 CAGGCCAACCTCATGGGATCAATGGGATCATGGGAAGTCTCATGGCTAACGCT 180  
 Qy 181 CTGGCGGGCCAACACCCCTCTGTTAGGGATCTGGGATGGGGCTCAAACAGGGTCACTC 240  
 Db 181 CTGGCGGGCCAACACCCCTCTGTTAGGGATGGGGCTCAAACAGGGTCACTC 240  
 Qy 241 ACACACCCATAACCAATTATCATGGATCATGGATCATGGATCATGGACTGAGGCTCTAC 300  
 Db 241 ACACACCCATAACCAATTATCATGGATCATGGATCATGGATCATGGACTGAGGCTCTAC 300  
 Qy 361 GGCAGGGGGCATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG 300  
 Db 361 GGCAGGGGGCATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG 300  
 Qy 421 AGGAGAGCTCTTACAGGGAGTCTGATAGATGGGGCTCTGGGGCTCTGGGGCTCTAC 480  
 Db 421 AGGAGAGCTCTTACAGGGAGTCTGATAGATGGGGCTCTGGGGCTCTGGGGCTCTAC 480  
 Qy 481 ATCGAGGAGGAATGAGCTGGCCAGTCAGGAAAGGGCTGGGGCTGGGGCTGGGGCTGGGG 540  
 Db 481 ATCGAGGAGGAATGAGCTGGCCAGTCAGGAAAGGGCTGGGGCTGGGGCTGGGGCTGGGG 540  
 Qy 541 ACAGCCCAAGCAAGGGAGGGAGCCCTGTCGGCTGGGGCTGGGGCTGGGGCTGGGG 600  
 Db 541 ACAGCCCAAGCAAGGGAGGGAGCCCTGTCGGCTGGGGCTGGGGCTGGGGCTGGGG 600  
 Qy 601 GAGACCTCTGGGAACACATGGGAACTCATGGGAACTCATGGGGATACAGTACTACGG 660  
 Db 601 GAGACCTCTGGGAACACATGGGAACTCATGGGAACTCATGGGGATACAGTACTACGG 660  
 Qy 661 TTGTCCTCTCTGGGAATCCGGGATGTCAGCTGTCAGGGTTACAGCCTCTGTC 720  
 Db 661 TTGTCCTCTCTGGGAATCCGGGATGTCAGCTGTCAGGGTTACAGCCTCTGTC 720  
 Qy 721 ACTAGCCGCTCACCAACCACTACCCCTCTGTTACACTCTCGGGGATGSGTAGCC 780  
 Db 721 ACTAGCCGCTCACCAACCACTACCCCTCTGTTACACTCTCGGGGATGSGTAGCC 780  
 Qy 781 GCCCAACTCTGCCCCCAGTGTCTCACTGCTTGAGGGCGGGATGCTGGTGG 840  
 Db 781 GCCCAACTCTGCCCCCAGTGTCTCACTGCTTGAGGGCGGGATGCTGGTGG 840  
 Qy 841 GCTGTTGCACTGGGAATGGGGCTATGGGA 900  
 Db 841 GCTGTTGCACTGGGAATGGGGCTATGGGA 900  
 Qy 901 GGAGTGCAGGGCGCTGTCGGCTTAAGTCATGAGCGCGAATGGCTCCACCGAG 960  
 Db 901 GGAGTGCAGGGCGCTGTCGGCTTAAGTCATGAGCGCGAATGGCTCCACCGAG 960  
 Qy 961 GACCTGGTTACTTACTCTGGCACTCTCTGTCGGCTGGGGATGCTGGTGG 1020  
 Db 961 GACCTGGTTACTTACTCTGGCACTCTCTGTCGGCTGGGGATGCTGGTGG 1020

QY 1021 TGGCCAGCGGAACTSGTCGCACTGGGTCAGGGGCTGGCAGTGGTGAC 1080  
 Db 1021 TGGCGACGCAACTGCGCTGGCACGTTGGTCAGGGGGCTGGCAGTGGTGAC 1080  
 QY 1081 CGGCTGATAGGTCTGCCTGGGGTACCATGTTCCCACCGACTATGGCAGAG 1140  
 Db 1081 CGGCTGATAGGTCTGCCTGGGGTACCATGTTCCCACCGACTATGGCAGAG 1140  
 QY 1141 AGCGAGCCGAGCACGCTGTAACATGGTCTCGACGCTACTATACCGAAGCT 1200  
 Db 1141 AGCGAGCCGAGCACGCTGTAACATGGTCTCGACGCTACTATACCGAAGCT 1200  
 QY 1261 AGGAGGTTGGACTGATGAGTTGGCTGACTCAAGCTTGCTGACTTCAGGAGCT 1320  
 Db 1261 AGGAGGTTGGACTGATGAGTTGGCTGACTTCAGGAGCT 1320  
 QY 1321 AAGCTCTGCGCGATTAACCGAGCTTCTCCACGCCCTGCGACTATGGCTGA 1380  
 Db 1321 AAGCTCTGCGCGATTAACCGAGCTTCTCCACGCCCTGCGACTATGGCTGA 1380  
 QY 1381 GCTCGGGGAGACCGCATATCGAGACCCCTGCTCATGGACACAGATACCGGA 1440  
 Db 1381 GCTCGGGGAGACCGCATATCGAGACCCCTGCTCATGGACACAGATACCGGA 1440  
 QY 1441 CATGCTCAAACGGTTCAGGATGTTGGCTPAGACTCTGGTAACTGGCAT 1500  
 Db 1441 CATGCTCAAACGGTTCAGGATGTTGGCTPAGACTCTGGTAACTGGCAT 1500  
 QY 1501 GGACATTCGGGCTGTGGGGCTGGCTGGCAT 1560  
 Db 1501 GGACATTCGGGCTGTGGGGCTGGCTGGCAT 1560  
 QY 1561 TATTCCAGGGGCTGTGGGGCTGGCTGGCAT 1620  
 Db 1561 TATTCCAGGGGCTGTGGGGCTGGCTGGCAT 1620  
 QY 1621 GATTCCACTCTGTGAGGAGTACAACGCTAACAGCTAAATGGCGTGCAGGGTCCA 1680  
 Db 1621 GATTCCACTCTGTGAGGAGTACAACGCTAACAGCTAAATGGCGTGCAGGGTCCA 1680  
 QY 1680 GACATTCGGGCTGTGGGGCTGGCTGGCAT 1740  
 Db 1680 GACATTCGGGCTGTGGGGCTGGCTGGCAT 1740  
 QY 1741 AACCTCTCTTCAAGAGTGGGGCTGGCTGGCAT 1800  
 Db 1741 AACCTCTCTTCAAGAGTGGGGCTGGCTGGCAT 1800  
 QY 1801 TGGCAGCTCCATGGGGGAGGAGTGGAGTGTGCTACTTCATGCTCAACG 1860  
 Db 1801 TGGCAGCTCCATGGGGGAGGAGTGGAGTGTGCTACTTCATGCTCAACG 1860  
 QY 1860 TGGCAGCTCCATGGGGGAGGAGTGGAGTGTGCTACTTCATGCTCAACG 1860  
 Db 1860 TGGCAGCTCCATGGGGGAGGAGTGGAGTGTGCTACTTCATGCTCAACG 1860  
 QY 1861 CCGTCCACATCACAGGAGGAGGCTAACATCCAGTCGGCTAACCATGTTGG 1920  
 Db 1861 CCGTCCACATCACAGGAGGAGGCTAACATCCAGTCGGCTAACCATGTTGG 1920  
 QY 1921 TGGCCAGCTCTTCAAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1980  
 Db 1921 TGGCCAGCTCTTCAAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1980  
 QY 1981 CAAATGACTTCCAGGCTGACTCATCGAGGCCAACCTCTGAGGCGCATGATG 2040  
 Db 1981 CAAATGACTTCCAGGCTGACTCATCGAGGCCAACCTCTGAGGCGCATGATG 2040  
 QY 2041 GGC 2043  
 Db 2041 GGC 2043

---

RESULT 3  
 US-09-324-977-35  
 Sequence 35, Application US/08324977  
 Patent No. 5747339  
 GENERAL INFORMATION:  
 APPLICANT: OKAYAMA, Hiroto  
 APPLICANT: FUKE, Isao  
 APPLICANT: MORI, Chisato  
 APPLICANT: TAKAMIZAWA, Akahisa  
 APPLICANT: YOSHIDA, Iwao  
 TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 CDNA AND ANTIGEN POLYPEPTIDE  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Armstrong, Websterman, Hattori, Mcieland &  
 STREET: 1725 K St. N.W. Suite 1000  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/324, 977  
 APPLICATION NUMBER: JP 2-230921  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/635, 451  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 18-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-305605  
 FILING DATE: 09-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099, 706  
 APPLICATION NUMBER: US 07/635, 451  
 FILING DATE: 31-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/769, 996  
 FILING DATE: 02-OCT-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/635, 451  
 FILING DATE: 28-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stevens-Smith, Theresa M.  
 REGISTRATION NUMBER: 36'281  
 REFERENCE/DOCKET NUMBER: 900703D  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEX: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 35:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7863 base pairs  
 FEATURE: CDS  
 NAME/KEY: CDS  
 LOCATION: 1..7863  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..7863  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..7863  
 OTHER INFORMATION: /note= "sequence = 1500 - 9362 of  
 OTHER INFORMATION: SEQ ID NO: 1"  
 US-08-324-977-35

Y	Matches	local Similarity	Pred.	No.	0;	M1 matches	154;	Indels	0;	Gaps	0;	Db	Qy
Y	1	TGGGAGGCGCTTCACAGGCCCTCACCCACGTGGATGCCACTTCTGTCCAAACAAG	60									4585	CGGCTGATATSGTTGCCTCGGSGSTAATCATGTTTCCCCAAGCACTATGTGCTGAG
b	3505	TGGGAGGAGTCTTACAGGCCCTACCCATAGTAGCTAGCAGACACTTCGTCCAGACAG	3564									4644	
b	61	CAGGAGGAGACAACCTTCCCTACTGTGTCGGTACCGACTACTGTGTCGGCTAGGCC	120									1141	ACGAGACCCGAGCAGGTGTCACTAGATCTCTCCAGCTTAATCACCCACTGTG
b	3565	CAGGAGGAGACAACCTTCCCTACTGTGTCGGCTAGGATACCGACAGGCTGAGGCC	3624									4704	AGGAGACCCGAGCAGGTGTCACTAGATCTCTCCAGCTTAATCACCCACTGTG
b	121	CAGGCCCACTCCATCTATGGATCAATGTGAAGTGTCTCATAGCTAACGCTAACCT	180									1201	AAGAGSCTCCACCGTGGATAAGAGGAGTGTCTCACCGCTTCACTACACTAACG
b	3625	CAGGCCCACTCCATCTATGGATCAATGTGAAGTGTCTCATAGCTAACGCTAACCT	3684									4705	AAAAGSCTCCACCGTGGATAAGAGGAGTGTCTCACCGCTTCACTACACTAACG
b	181	CTGGCGGGGCCAACACCCCTGTGTAGGTGGAGGCCGTCACAAACAGAGGTCAACCTC	240									1261	AGGAGATTTGGACTGTGATGCACAGTTGGCTGACTCAAGACCTGCTCCAGTCC
b	3685	CTGCAAGGGCAACACCCCTGTGTAGGTGGAGGCCGTCACAAACAGAGGTCAACCTC	3744									4765	AGGAGATTTGGACTGTGATGCACAGTTGGCTGACTCAAGACCTGCTCCAGTCC
b	241	ACACCCCAATAACCAAATCATCATGTGTCAGTCAGCGACACTGGAGCTCCTACG	300									1321	AAGCTCTGCGGCCGATTACCGGGAGTCCAGGCTGAGCTGGAGCTCCTACG
b	3745	ACCCACCCATAACCAAATCATCATGTGTCAGTCAGCGACCTGGCTGACCTGGAGT	3804									4825	ANGCTCTGCGGCCGATTACCGGGAGTCCAGGCTGAGCTGGCTGACCTGGAGT
b	301	AGCACCTGGGTCTGGGGGGCTCTGGAGCTCTGGGTGCTGTTGACACA	360									1381	GTCTGGCGGCCGAGACGCGCTACCTGGAGTCCAGGCTGAGCTGGCTGACACA
b	3805	AGCACCTGGGTCTGGGGGGCTCTGGAGCTCTGGGTGCTGTTGACACA	3864									4885	GTCTGGCGGCCGAGACGCGCTACCTGGAGTCCAGGCTGAGCTGGCTGACACA
b	361	GGCAGCGTGGTCATGTGGTAGGATGATCTGTGTCGGGCCGCTATGTTCCGAC	420									1441	CATGCTAAACGGTCCATGAGTGGCTAAGACCTGCTGACTCAAGACCTGCTCC
b	3865	GGCAGCTGTGGTCATGTGGTAGGATGATCTGTGTCGGAGGCCGCTATGTTCCGAC	3924									4945	CATGCTAAACGGTCCATGAGTGGCTGACTCAAGACCTGCTCCAGTCCAGTCC
b	421	AGGAAAGTCTTACAGGAGTCTGAGTGGAGACTGGCGCTGCACTCCCTAC	480									5004	1501 GGACATTCCCATCAACGCTACACGCTACATGAGCTGAGCACCTCTCATG
b	3925	AGGGAGCTCTACAGGAGTCTGATGAAATGGAGAGTGCCTGACCTCCCTAC	3984									5005	5005 GGACATTCCCATCAACGCTACACGCTACATGAGCTGAGCACCTCTCATG
b	481	ATCGAGGAGGATGGCTGGCGAGCAGTCAAGAAAGCGCTGGGTTGCTGAG	540									1561 TATCCAGGGGGCTGGCGCGTGTGAGGAGTACGTGGAGTTACCGGGTGGGG	
b	3985	ATCGAGGAGGATGGCTGGCGAGCAGTCAAGAAAGCGCTGGGTTGCTGAG	4044									5065 TATTCAGGGGGCTGGCGCGTGTGAGGAGTACGTGGAGTTACCGGGTGGGG	
b	541	ACAGGCCAACAGGGAGGCTGTCGGCTGCTCCGGTGGAGTCCAGTGGAGCC	600									1621 GATTCCACTACGTGACGAGGATGACACTGACAGCTGAAATGGCGTCCAG	
b	4045	ACAGGCCAACAGGGAGGCTGTCGGCTGCTCCGGTGGAGTCCAGTGGAGCC	4104									5125 GATTCCACTACGTGACGAGGATGACACTGACAGCTGAAATGGCGTCCAG	
b	601	GAGGCTCTGGGCAACACATGTCGAACTCATGAGCGGAGTACAGTACTACGGC	660									1681 GCCCGCGAATCTTCACAGAAGTGGATGGATGGGGGTCGCGTCAAGGTA	
b	4105	GAGACATTCTGGGCAACAGCATGTCGAAATTTCATGACGCGGATACAGTACGGC	4164									5185 GCTCTGAAATTCTCTGGAGGTTGAGCGGACTGCGGGTGCACGGTAGCT	
b	61	TGTGCACTCTGGGAAATCCGGGATTCATCATGATGGCGTTCACGCCCTGTC	720									1741 AACCTCTCTGGAGGGGTCACATCCAGTGGGGGTCACACAAATACGTGGGG	
b	4165	TTATCCACTCTGGGAAACCCGCAATGACATCTGTGATGGATTCACGCCCTATC	4224									5245 AGGGCTCTCTGGAGGGGAGGTTACATCCAGTGGGGGTCACACAAATACGTGGGG	
b	721	ACTAGCCCTCACCCCAATCTACCTCTGTGTCAGTACATCTGGGGATGGTAGCC	780									1801 TCGCAGCTCCCAGTGAGGGGACCGAGATGTCAGCTGCTCACCTCACCGAC	
b	4225	ACCAAGCCCGCTACACCAACAGTACCCCTCTGGGTTAACATCTGGGGGGGGCT	4284									5305 TCACAGTGTACCGTGCAGGCCGACCGGAGTGTGAGTGCAGTGCACCTCACCGAC	
b	781	GCCCAACTCTCCCCCACTGTGTCAGTACCTGGAGGGGCTGAGCGGAGTGTGCG	840									1861 CCCTCCACATCACAGAGAGGGCTGAGCGGAGTGTGAGTGCACCTCACCGAC	
b	4285	GCCCAACTCTCCCCCACTGTGTCAGTACCTGGAGGGGCTGAGCGGAGTGTGCG	4344									5365 CCCTCCACATCACAGAGAGGGCTGAGCGGAGTGTGAGTGCACCTCACCGAC	
b	841	GCTGTGCGAGCATAGGCCCTGGGAAGGTGTCAGACATCTGGGGGTATGGCA	900									1921 TGGCCAGCTCTTCAGTGTGAGTGCAGTGCCTGCGGCTTCCTGAGGAGCATAC	
b	4345	GCTGTGCGAGCATAGGCCCTGGGAAGGTGTCAGACATCTGGGGGTATGGCA	4404									5425 TGGCCAGCTCTTCAGTGTGAGTGCAGTGCCTGCGGCTTCCTGAGGAGCATAC	
b	901	GGAGTGGCGAGGCCCTGTGGCTTAAAGTCATGAGCGCGAATGCTTCACCGAG	960									1981 CAAATGACTCCAGAGCTCACCTCATGAGGCCACCTCTGGCGGCATGAGATG	
b	4405	GGAGTGGCGAGGCCCTGTGGCTTAAAGTCATGAGCGCGAATGCTTCACCGAG	4464									5454 CACAGTGTCTCGGAGCTGACCTCATCGAGGCCAACCTCTGGCGGGAGAGT	
b	961	GACCTGTGAGTACTTCTCCCTGTGGCTTAAAGTCATGAGCGCGAATGCTTCACCGAG	1020									2041 GGC 2043	
b	4465	GACCTGTGAGTACTTCTCCCTGTGGCTTAAAGTCATGAGCGCGAATGCTTCACCGAG	4524									5545 GGC 5547	
Y	1021	TGGCGAGGAGTACTGGCTGGGACCTGGTGGCTAGGGAGGGGCTGTCAGTGGATGAC	1080									RESULT 4	
Y												US-08-384-616-35	





QY	1141	AGCGAGCCGCGAGCAGGGTCACTCGAGATCTCTCGAGCTTA	1200
Db	1261	AGGGATGTTGGAGACTTGATGCAAGTTTGCTGACTGCTG	4704
Db	4645	AGCAGACGCCAGCGCGTGTACTCGAGTCCTCTCGAGCTTAC	4704
QY	1201	AAGAGCTCCACCAAGCTGGATAACAGGAAGCTGCTCCAGC	1260
Db	4705	AAAGCCTCCACCACTGGATTAGTACAGTGCCTCCACCGT	4764
QY	1321	AAGCTCTGCGGATACCGGAGTCCCTTTCTCATGCGAAC	1380
Db	4825	AAGCTCTGCGGAGACTACCTGGAGTCCCTTTCTCGTGC	4884
QY	1381	GTCCTGGGGAGAGCGCATCTGCAACCTCCATGAGGAG	1380
Db	4765	AGGGATGTTGGAGCTGATGACGGTGTGACTGACTGAC	4824
QY	1431	AAGCTCTGCGGATACCGGAGTCCCTTTCTCATGCGAAC	1440
Db	4885	GTCCTGGGGAGAGCGCATCTGCAACCTCCATGAGGAG	1440
QY	1441	CATGCTAAACAGGTTCTAGGAGTAGTGTGGGCTTA	1500
Db	4945	CATGCTAAACAGGTTCTAGGAGTAGTGTGGGCTTA	5004
QY	1501	GGAACTTCCCATAACGATACAGCAACACACCTCC	1560
Db	5005	GGAACTTCCCATAACGATACAGCAACACACCTCC	5064
QY	1561	TATCCAGGGCTGTGGGGTGCACACGGCCCTTGAC	1620
Db	5065	TATCCAGGGCTGTGGGGTGCACACGGCCCTTGAC	5124
QY	1621	GATTCACTAAGGAGCATGACACTGAGGAGTAAAT	1680
Db	5125	GATTCACTAAGGAGCATGACACTGAGGAGTAAAT	5184
QY	1681	GCCCCGAATCTTCAAGAGTGGTGGCTGACAGTGT	1740
Db	5185	GCTCTCGAATCTTCAAGAGTGGTGGCTGACAGTGT	5244
QY	1741	AAACCTCTTACGGAGGAGTCACATTCAAGGTTGGCT	1800
Db	5245	AAACCTCTTACGGAGGAGTCACATTCAAGGTTGGCT	5304
QY	1801	TGGCACTCCATGGAGCCGAACCGGAGTAGGTTCTAC	1860
Db	5305	TCACTACATCATGGAGCCGAACCGGAGTAGGTTCTAC	5364
QY	1861	CCCTCCACATCACACAGAACGGTAAGGGAGCTGGC	1920
Db	5365	CCCTCCACATCACACAGAACGGTAAGGGAGCTGGC	5424
QY	1921	TTGGCCAGCTTCACTAACGGTTCTCGCCGCTTCCTG	1980
Db	5425	TTGGCCAGCTTCACTAACGGTTCTCGCCGCTTCCTG	5484
QY	1981	CAAATGACTTCCAGACGCTGACCTCATCGAGGCCAA	2040
Db	5485	CACCATGCTCTCCGACGGTGTACCTCATGAGGCAAC	5544
QY	2041	GGC 2043	
Db	5545	GGC 5547	



Db 4585 CGGCTGATAGCTTGCCTTGCGGGGTATCATTTCCCCACGACTATGTCGAG 4644  
 Qy 1141 AGCGAACSGCGACGACGCTACTAGATCTCTCGACCTTAATACCCAACTGTG 1200  
 Db 4645 AGCGAACSGCGACGCGTGTACTAGATCTCTCGACCTTAATACCCAACTGTG 4704  
 Qy 1261 AGGGATGTTGGGACTGGATATGCAAGGGACTGCCAGTTCACCCCTCCTCGGCTA 1260  
 Db 4765 AGGGATGTTGGGACTGGATATGCAAGGGACTGCCAGTTCACCCCTCCTCGGCTA 4764  
 Qy 1321 AGGCTCTGCGGGATTAACGGGAATGCCAGTTCACCCCTTCTCATGCCAGTCGCTCACCGGTTCCGGCTA 4760  
 Db 4825 AGGCTCTGCGGGATTAACGGGAATGCCAGTTCACCCCTTCTCATGCCAGTCGCTCACCGGTTCCGGCTA 4824  
 Qy 1381 GTCGCGGGAGAGGGCATACCGACACCTGCTCATGGACACAGTACCGGA 1440  
 Db 4885 GTCGCGGGAGAGGGATCATGCCAACCCATGCTGGAGACAGATACCGGA 4944  
 Qy 1441 CATGTCAAAAACGTTCCATGAGGATCGTGGCCCTAACGACTTGAGTACATGGCAT 1500  
 Db 4945 CATGTCAAAAACGTTCCATGAGGATCGTGGCCCTAACGACTTGAGTACATGGCAT 5004  
 Qy 1501 GGAACTATCCCATAAGCATACACCAAGGGCCCTGAGCCTCCAGGCCAAC 1560  
 Db 5005 GGAACTATCCCATAACCATACACCAAGGGCCCTGAGCAGACCTGCGAAC 5064  
 Qy 1561 TATTCAGGGCTGTGGGGGGCTGTGAGGAGTACTGAGGTTACGGGGGGGG 1620  
 Db 5065 TAATCTAGGGCTGTGGGGGGCTGTGAGGAGTACTGAGGTTACGGGGGG 5124  
 Qy 1621 GATTCTCACTACGGTACGAGCATGACCACTGACAGGTTAAATCCCGTCCAGGTCCA 1680  
 Db 5125 GATTCTCACTACGGTACGAGCATGACCACTGACAGGTTAAATCCCGTCCAGGTCCA 5184  
 Qy 1681 GCCCCGGAATCTTCACAGAAGTGAATGGGGTGGGTGACAGGTACGGTCC 1740  
 Db 5185 GTCCTGAAATCTTCGGGGTGGACGGAGTGGGGTGGACAGGTACGGTCC 5244  
 Qy 1741 AACCTCTTACGGAGGGTCACATTCAGGTGGGCTCAACCAATACCTGGTGG 1800  
 Db 5245 AGGCTCTCTACGGGGAGGGTTACATTCAGGTGGGCTCAACCAATACCTGGTGG 5304  
 1801 TCGGCTCCATGGAGGCCAACGGATGTTAGGAGTGTCAATTCCATGTCACCGAC 1860  
 Db 5305 TCACTGCTACATGGAGGCCAACGGATGTTAGGAGTGTCAATTCCATGTCACCGAC 5364  
 Db 5364 TCACTGCTACATGGAGGCCAACGGATGTTAGGAGTGTCAATTCCATGTCACCGAC 5364  
 Qy 1861 CCCCTCCACATCACGGCAGACGCCCTAACGCCAGCTGCGCAGGGGTCTCCCTC 1920  
 Db 5365 CCCCTCCACATCACGGCAGACGCCCTAACGCCAGTGGCTGGCCAGGGGTCTCCCTC 5424  
 Qy 1921 TTGGCAGCTCTTCAGCTAGCCAGTGTCTGGCTTCTCGAAGGGACATACATTAC 1980  
 Db 5425 TTGGCAGCTCTTCAGCTAGCCAGTGTCTGGCTTCTCGAAGGGACATACATTAC 5484  
 Qy 1981 CAAMATGACTTCCGAGGCTGACTCATGAGCCAACTCTCTGGGGGCTAGATG 2040  
 Db 5485 CACCATGCTCTCCGGACGCTGACTCATGAGCCAACTCTCTGGGGGCTAGATG 5544  
 Qy 2041 GGC 2043  
 Db 5545 GGC 5547

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/324,977  
 FILING DATE: 18-OCT-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 31-AUG-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-305605  
 FILING DATE: 09-NOV-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099,706  
 FILING DATE: 30-JUL-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/769,996  
 FILING DATE: 02-OCT-1991  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/635,451  
 FILING DATE: 28-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stevens-Smith, Theresa M.  
 REGISTRATION NUMBER: 36,251  
 REFERENCE/DOCKET NUMBER: 900703D  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 TELEX: 440142  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7917 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE: NAME/KEY: misc feature  
 LOCATION: 1..7916  
 OTHER INFORMATION: /note= "sequence = 1500 - 9416 of  
 OTHER INFORMATION: SEQ ID NO: 1"  
 US-08-324-977-31

116-08-334-977-3

Sequence 31,

## **; GENERAL INFORMATION**

Db 3505 TGGAGAGTCCTCACAGGCTCACCATATAGTGCACACTCTGTCCAGACAG 3564  
 Qy 61 CAGGAGAGAACACTTCCCTACCTGTTGAGTACCGGTTACAGGTTACTGTTG 120  
 Db 3565 CAGGAGAGAACACTTCCCTACCTGTTGAGTACCGGTTACAGGTTACTGTTG 3624  
 Qy 121 CNGGCCCACTTCATCATGGTAATGTTGAAAGTCTCATACGGTAAGCCACT 180  
 Db 3625 CAGGCCCACTTCATCATGGTAATGTTGAAAGTCTCATACGGTAAGCCACT 3684  
 Qy 181 CTGGCGGGCAACACCTTGTTGATAAGGCTGGAGGCGTCAAACAGGAACTC 240  
 Db 3685 CTGGCGGGCAACACCTTGTTGATAAGGCTGGAGGCGTCAAACAGGAACTC 3744  
 Qy 241 ACACACCCATACCAAAATTCTATGTCATGTCATGTCAGGCTGGAGGCG 300  
 Db 3745 ACCACCCATACCAAAATCATCATGTCATGTCAGGCTGGAGGCGTCAAAC 3804  
 Qy 301 AGCACCTGGTCTGGTGGGGCTTSCAGCTCTGGCGGTTATGGCTGACAA 360  
 Db 3805 AGCACCTGGTCTGGTGGGGCTTSCAGCTCTGGCGGTTATGGCTGACAA 3864  
 Qy 361 GCGAGCTGGTCTATGGGGTAGATCATCTGTCCGGG36CCG3TATTCGCGAC 420  
 Db 3865 GCGAGCTGGTCTATGGGGTAGATCATCTGTCCGGG36CCG3TATTCGCGAC 3924  
 Qy 421 AGGGAAGCTCTTACCAAGGATGATGAGTGGAGAGTGGGGCTGGCTGCG 480  
 Db 3925 AGGGAGCTCTTACCAAGGATGATGAGTGGAGAGTGGGGCTGGCTGCG 3984  
 Qy 481 ATGAGSCAGGAAATGCGTCCCGCGACTTCAGTGGAGAGTGGGGCTGGCG 540  
 Db 3985 ATGAGSCAGGAAATGCGTCCCGCGACTTCAGTGGAGAGTGGGGCTGGCG 4044  
 Qy 541 AGAGCACCAGCAAGGGGAGCCGCTGCTCCGGTGGAGTCCAAAGTGGAGCCT 600  
 Db 4045 AGAGCACCAGCAAGGGGAGCCGCTGCTCCGGTGGAGTCCAAAGTGGAGCCT 4104  
 Qy 601 GAGACCTTCCTGGAAACACATGTAACATCAGCGGAACTAGTACTTACGGC 660  
 Db 4105 GAGACATTCTGGCGAAGCAATGTCAGCGGAACTAGTACTTACGGC 4164  
 Qy 661 TTGTCCTCTCTGGAAACACATGTAACATCAGCGGAACTAGTACTTACGGC 720  
 Db 4165 TTATCCACTCTCTGGAAACACCGCAATAGTACTTACGGC 4224  
 Qy 721 ACTAGCGCTTACACCAACTTACCCCTCTGCTAACATCTGGGGATGGTAGCC 780  
 Db 4225 AccAGCGCTTACACCAACTTACCCCTCTGCTAACATCTGGGGATGGTAGCC 4284  
 Qy 781 GCCCAACTCGTCTCCCAGTCTGCTTCACTTCTAGCGCCCATCTGGGG 840  
 Db 4285 GCCCAACTCGCCCCCAGGCCGCTTCGGGGCGCCGATGCCGGTGGCG 4344  
 Qy 841 GCTGTTGAGACATAGCGCTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 900  
 Db 4345 GCTGTTGAGACATAGCGCTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 4404  
 Qy 901 GAGTGGAGGAGGGCTGTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 960  
 Db 4405 GAGTGGAGGAGGGCTGTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 4464  
 Qy 961 GACCTGTTAATCTACCTCGCTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 1020  
 Db 4465 GACCTGTTAATCTACCTCGCTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 4524  
 Qy 1021 TGGAGAGTCCTCACAGGCTCACCATATAGTGCACACTCTGTCCAGACAG 1080  
 Db 4525 TGGAGAGTCCTCACAGGCTGGAGGCGTGGAGGCGTGGAGGCGTGGAGGAC 4584  
 Qy 1081 CGCTGTAGAGTGTGCTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 1140  
 Db 4585 CGCTGTAGAGTGTGCTTGGAGAGTGTGCTGACACCTGGGGGATGGAGCA 4644

RESULT 8  
 US-08-384-616-31  
 ; Sequence 31, Application US/08384616  
 ; Patent No. 5,847,01  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OKAYAMA, Hiroto  
 ; APPLICANT: FUKE, Isao

APPLICANT: MORI, Chisato  
 APPLICANT: TAKAMIZAWA, Akahisa  
 APPLICANT: YOSHIDA, Iwao  
 TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSE: Armstrong, Westerman, Hattori, McLeland &  
 STREET: 1725 K ST. N.W. Suite 1000  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Disquette, 3.5 in, 1.44MB  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/384,616  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/169,996  
 FILING DATE: 02-OCT-1991  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 31-AUG-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stevens-Smitch, Theresa M.  
 REGISTRATION NUMBER: 36,281  
 REFERENCE/DOCKET NUMBER: 900703B  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-3930  
 TELEX: 440142  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7917 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..7862  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1..7916  
 OTHER INFORMATION: /note= "sequence = 1500 - 9416 of  
 OTHER INFORMATION: SEQ ID NO: 1"  
 US-08-384-616-31

Query Match Similarity 87.9%; Score 1796.6; DB 2; Length 7917;  
 Best Local Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 1 TGGAGGAGTCCTCCACAGGCCAACCGACGATGATGCCACTTCTGGCTCAAAAG 60  
 Db 3505 TGGAGGAGTCCTCCACAGGCCAACCGACGATGATGCCACTTCTGGCTCAAAAG 3564  
 Qy 61 CAGGAGGAGAACCTCCACCTGGTGTACCTGGTGTACCGGTTACTGTTGGCTAGGGC 120  
 Db 3565 CAGGAGGAGAACCTCCACCTGGTGTACCTGGTGTACCGGTTACTGTTGGCTAGGGC 3624

Qy 121 CAGGCCAACCTCCATCTGGATAATTGGAGTGCTCATGGCTTAAGGCTACT 180  
 Db 3625 CAGGCCAACCTCCATCTGGATCAATTGGAGTGCTCATGGCTTAAGGCTACT 3684  
 Qy 181 CTGGGGGGCAACACCCCTGGTGTAGCTGGCTGGAGGCCTCCAAAAGGTTACCTC 240  
 Db 3685 CTGAGGGSACACCCCTGGTGTAGCTGGCTGGAGGCCTCCAAAAGGTTACCTC 3744  
 Qy 241 ACACACCCATAACCAATTCTATCATGCGATGCGATGCGATGCGATGCG 300  
 Db 3745 ACCACCCATACACATGCGATGCGATGCGATGCGATGCGATGCGATGCG 3804  
 Qy 301 AGCACCTGGCTGGCTGGAGGCTCTGGCTGGCTGGCTGGCTGGCTGGCTGG 360  
 Db 3805 AGCACCTGGCTGGCTGGAGGCTCTGGCTGGCTGGCTGGCTGGCTGG 3864  
 Qy 361 GGCAGGGGGTGTGGTAGGTCATCTGGCGGGGGCTATGGTCCCAC 420  
 Db 3865 GGACTGTGTCATGGTAGGTCATCTGGCGGGGGCTATGGTCCCAC 3924  
 Qy 421 AGGAAGTCTCTACAGGAGTTCATGAGATGATGATGAGATGAGAGTGG 480  
 Db 3925 AGGAGCTCTACAGGAGTTCATGAGATGATGAGATGAGAGTGG 3984  
 Qy 481 ATCGAGGAGGAATGCGAGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 540  
 Db 3985 ATCGAGGAGGAATGCGAGCTGGCTGGCTGGCTGGCTGGCTGG 4044  
 Qy 541 ACAGCCACAGGAGGGGGCTGTCCTGGCTGGCTGGAGTCAGGGAGGCC 600  
 Db 4045 ACAGCCACACAAAGGGAGGCTGTCCTGGCTGGCTGGCTGGCTGG 4104  
 Qy 601 GAGACCTCTGGGAAACATGTCATCGGGATACTGACTTAGGGC 660  
 Db 4105 GAGACCTCTGGGAAACATGTCATCGGGATACTGAGGGATACTGAGGC 4164  
 Qy 661 TTGTCACATCTGGCTGGAAATCCCGGATGTCATCATGTCATGTC 720  
 Db 4165 TTATCCACTCTGCCTGGANACCCGATAGCATCATGTCATGTCATGTC 4224  
 Qy 721 ACTACCCGTCACACCCATCTCCCTCTGGCTAACATCCCTGGGGATGG 780  
 Db 4225 ACCACCCGCTAACCCCAAGTACCTCTGGCTGGCTGGCTGGCTGG 4284  
 Qy 781 GCCCAAATCGTCCTCCCCAGTGTCTGTTAGCTTGTAGGGGGATGGC 840  
 Db 4285 GCCCAAATCGCCCTTACGGCTTGCGCTTGCGCTTGCGCTTGCG 4344  
 Qy 841 GCTGTGCGACATAGGCCCTGGAGAGTCTTGTGGACATCTGGGGCTATGGCA 900  
 Db 4345 GCTGTGCGACATAGGCCCTGGAGAGTCTTGTGGACATCTGGGGTTATGGCA 4404  
 Qy 901 CGAGTGGCAGGGGCTTGGGGCTTAAGTCATGAGGGGGAAATGCCCTCCACGG 960  
 Db 4405 GGAGTGGCAGGGGCTTGGGGCTTAAGTCATGAGGGGGAAATGCCCTCCACGG 4464  
 Qy 961 GACCTGGTAACTACTCCCTGCACTCTCTCTGGCTGGCTTAAGTCATGAGGGGGCTATGG 1020  
 Db 4465 GACCTGGTAACTACTCCCTGCACTCTCTGGCTGGCTGGCTTAAGTCATGAGGGGGCTATGG 4524  
 Qy 1021 TCGGGAGGATACTGGCTGGCAGGGGGCTGGCTGGAGGGGGCTGGCTGG 1080  
 Db 4525 TGTGAGCAATACTGGCTGGCTGGAGGGGGCTGGCTGGAGGGGGCTGG 4584  
 Qy 1081 CGGTGATAGCTGGCTGGGGGTAACTGTTCCCGCCACCATATGGCCAG 1140  
 Db 4585 CGGTGATAGCTGGCTGGAGGGGGCTGGCTGGAGGGGGCTGGCTGG 4644  
 Qy 1141 AGCGAGGCCGAGCACTGGTGTACCTGGTGTACCTGGCTGGCTGG 1200  
 Db 4645 AGCGAGGCCGAGGGGGTGTACCTGGCTGGCTGGCTGGCTGG 4704

QY 1201 AAGAGGCTCCACCAACTGGGTTAACCGAGACTGCTCAGCCCTTCCTCGCGCTCTGGCTA 1260

Db 4705 AAAGGCTCCACCAACTGGGTTAACCGAGACTGCTCAGCCCTTCCTCGCGCTCTGGCTA 4764

QY 1261 AGGGAGTGTGGGAATGGGATATGGGAACTGGGATATGGGAACTGGGATCTGACTTCAGACCTTCAGCT 1320

Db 4765 AGGGATGTTGGAACTGGATATGGGAACTGGGATCTGACTTCAGACCTTCAGCT 4824

QY 1321 AAGCTTCTTCCGGGAACTACGGGAGTCCCTTTCATGCCAACGTTGAGAG 1380

Db 4825 AAGCTTCTTCCGGGAACTACGGGAGTCCCTTTCATGCCAACGTTGAGAG 4884

QY 1381 GTCCTGCGGGAGAGGGCATCTGGATCTGGCTCTTTCATGCCAACGTTGAGAG 1440

Db 4885 GTCCTGCGGGAGAGGGCATCTGGATCTGGCTCTTTCATGCCAACGTTGAGAG 4944

QY 1441 CATGCTAAACACGGTTCATGGGGAGCTGAGGACCTGCTCATGGAGACAGATCACCGA 1500

Db 4945 CATGCTAAACACGGTTCATGGGGAGCTGAGGACCTGCTCATGGAGACAGATCACCGA 5004

QY 1501 GGAAACATTCCCCATACACCATACCCACGGGCCCTGACGGCCCTCCAGGCGAAC 1560

Db 5005 GGAAACATTCCCCATACACCATACCCACGGGCCCTGACGGCCCTGACACCCCTCAGGGCANCAC 5064

QY 1561 TATTCCTGAGGCGCTTGGGGGCTGGCTGTTGAGGAGTAAGCTGGGAT 1620

Db 5065 TATTCCTGAGGCGCTTGGGGGCTGGCTGTTGAGGAGTAAGCTGGGAT 5124

QY 1621 GATTTCCACTAGTGTGGGAGCATGGCCACTGACAGGTTAACCGGGTAAATCCCGTGCCAGGTCA 1680

Db 5125 GATTTCCACTAGTGTGGGAGCATGGCCACTGACAGGTTAACCGGGTAAATCCCGTGCCAGGTCA 5184

QY 1681 GCCCCCGAAATCTTCACAGAGTGGATGGGGTGGCGCTCACAGGTGGCTCCAGGTTACGGCTGG 1740

Db 5185 GCTCTGTGAACTCTCTCGGGAGGTGGACGGAGTGGCTTCAACAGGTACGCTCCGGTGC 5244

QY 1741 AACCTCTTACCGGGAGGGTCACTTCAGGCTGACGGCATGACCAATACCTGGTGG 1800

Db 5245 AGGCTCTCTACCGGGAGGGTCACTTCAGGCTGACGGCATGACCAATACCTGGTGG 5304

Db 1801 TCGAGCTCCATGGAGCCGAAACGGGATGACGGACTCTACTCCATGTCACCGAC 1860

QY 5305 TCAAGCTTACCATGGAGCCGAAACGGGATGACGGACTCTACTCCATGTCACCGAC 5364

Db 1861 CCCCTCCCACATCAGCAGCAGACGGCTAACGGCAGCTGGCCAGGGGTCTCCCTCC 1920

Db 5365 CCTCTCCCACATCAGCAGCAGACGGCTAACGGCAGCTGGCCAGGGTCTCCCTCC 5424

QY 1921 TTGGCAGCTCTTCAGCTAGCCAGTGTGCTGGCCCTCTTGAGGGCATGAGTG 1980

Db 5425 TTGGCAGCTCTTCAGCTAGCCAGTGTGCTGGCCCTCTTGAGGGCATGAGTG 5484

QY 1981 CAATGACTTCCAGACCTGACCTCATGGGCCAACCTCTGGGGCATGAGTG 2040

Db 5485 CACCATGCTCTGGAGCTGACCTCATGGGCCAACCTCTGGGGCATGAGTG 5544

QY 2041 GGC 2043

Db 5545 GGC 5547

RESULT 9

US-08-904-666A-31

; Sequence 31, Application US/08904666A.

; Patent No. 598130

; GENERAL INFORMATION:

; APPLICANT: OKAYAMA, Hiroto

; APPLICANT: FUKE, Iwao

; APPLICANT: MORI, Chiato

; APPLICANT: TAKAMIZAWA, Akahisa

; APPLICANT: YOSHIDA, Iwao

TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC

TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: Armstrong, Westerman, Hattori, McLeland & Naughton

STREET: 1725 K St. N.W. Suite 1000

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20006

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0

SOFTWARE: ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/904,686A

PRIOR APPLICATION DATA:

FILING DATE: 01-AUG-1997

PRIOR APPLICATION NUMBER: US 08/324,977

PRIOR APPLICATION DATA:

FILING DATE: 18-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-167466

FILING DATE: 25-JUN-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-230921

FILING DATE: 31-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 2-305605

FILING DATE: 09-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/039,705

FILING DATE: 30-JUL-1993

PRIOR ATTORNEY/ AGENT INFORMATION:

NAME: Michael and Lee-Nhung

REGISTRATION NUMBER: 31,541

REFERENCE DOCKET NUMBER: 90703G

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 659-2330

TELEFAX: (202) 887-0357

INFORMATION FOR SEQ ID NO: 31:

SEQUENCE CHARACTERISTICS:

LENGTH: 7917 base pairs

FEATURE: nucleic acid

FEATURE: misc feature

NAME/KEY: CDS

LOCATION: 1..7862

FEATURE: strandness

STRANEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA from genomic RNA

NAME/KEY: misc feature

LOCATION: 1..7916

LOCATION: 1..7916

OTHER INFORMATION: /note= "Sequence = 1500 - 9416 of

OTHER INFORMATION: SEQ ID NO: 1"

US-08-904-666A-31

Query Match Best Local Similarity Score 1796.6; DB 2; Length 7917; Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 1 TGGAGGAGCTTCATGGCTTACGGCTACCCACGGGATGCCACTTCCTGTCCTAACAAAG 60

Db 3505 TGGAGGAGCTTCATGGCTTACGGCTACCCACGGGATGCCACTTCCTGTCCTAACAAAG 3564

QY 61 CGGGAGGAGAACCTCCCTTACCGGGTACAGGTACTCTGTCCTAGGG 120

Db 3565 CGGGAGGAGAACCTCCCTTACCGGGTACAGGTACTCTGTCCTAGGG 3624

QY 121 CAGGCCCCACCTCCATCATGGGATCAATGGGAAGTGTCTACGGCTAACGCTACT 180  
Db 3625 CAGGCCCCACCTCCATCATGGGATCAATGGGAAGTGTCTACGGCTAACGCTACT 3684  
QY 181 CTGGCGGGGCCCCACACCTTGTGTTAGGTGGGGACGGCTCAAACGAGGTACCCCTC 240  
Db 3685 CTGGCACGGGCCCCACACCTTGTGTTAGGTGGGGACGGCTCAAACGAGGTACCCCTC 3744  
QY 241 ACCACACCCATTAACCAAAATTATCATGGGATCATGTCAGCTGGGAGCTGGGTG 300  
Db 3745 ACCACACCCATTAACCAAAATTATCATGGGATCATGTCAGCTGGGAGCTGGGTG 3804  
QY 301 AGCACCTGGGCTGGGGGGGCTTGAGCTCTGGGGGTTGAGCTGGGGGTTGACA 360  
Db 3805 AGCACCTGGGCTGGGGGCTTGAGCTGGGGGTTGAGCTGGGGGTTGACA 3864  
QY 361 GGCAGCTGGGATATGGGGAGGATCATCTGTCGGGGGGGCTGGGGGTTGACA 420  
Db 3865 GGCAGCTGGGATATGGGGAGGATCATCTGTCGGGGGGGCTGGGGGTTGACA 3924  
QY 421 AGGGAAGTCTCTTACCGGAACTCGATGAGATGAGCTGGGGGGGCTGGGGGTT 480  
Db 3925 AGGGAGCTCTTACCGGAACTCGATGAGATGAGCTGGGGGGGCTGGGGGTT 3984  
QY 481 ATCGAGGAGGAAATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 540  
Db 3985 ATCGAGGAGGAAATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4044  
QY 541 AGCGCCACCAACGAGGGGAACGGGGCTGGCTCCGGTGGGGGGGGGGGGGG 600  
Db 4045 AGCGCCACCAACGAGGGGAACGGGGCTGGCTCCGGTGGGGGGGGGGGGGG 4104  
QY 601 GAGACCTCTGGCGGAAACAGTGGAAACTCATCGGGGATACAGTAGCTAGGGC 660  
Db 4105 GAGACATCTGGCGGAACTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4164  
QY 661 TATGCCACTCTGGGAATCCGGGATTGATCACTGATGGGTTGACGGCTCTGTC 720  
Db 4165 TTATCCACTCTGGGAACCCGGAAATGCACTCATGATGGCATTCAGCTCTATC 4224  
QY 721 ACTAGCGGCTTACCAACCAATCTACCCCTCTGGTAACTACAGTGGGGGGGGGG 780  
Db 4225 ACCAGCGGCTTACCAACCAATCTACCCCTCTGGTAACTACAGTGGGGGGGGGG 4284  
QY 781 GGCACACTGGCTTCCCGGAACTGGCTGGCTTACAGCTGGGGGGGGGGGGGG 840  
Db 4285 GGCACACTGGCTTCCCGGAACTGGCTGGCTTACAGCTGGGGGGGGGGGGGG 4344  
QY 841 GCGTGGCGAGATAGGCTTGGGAAGGTGTGTTGGACATCTGGGGGGGGGGGG 900  
Db 4345 GCGTGGCGAGATAGGCTTGGGAAGGTGTGTTGGACATCTGGGGGGGGGGGG 4404  
QY 901 GGAGTGGAGGGGGCTGGGGCTTAAAGGTCATGGGGGAAATGCCCTCACCGAG 960  
Db 4405 GGAGTGGAGGGGGCTGGGGCTTAAAGGTCATGGGGGAAATGCCCTCACCGAG 4464  
QY 961 GACCTGGTCACTCACTCTGGCATCTCTCTGGGGCTGGGGCTGGGGCTGG 1020  
Db 4465 GACCTGGTCACTCACTCTGGCATCTCTCTGGGGCTGGGGCTGGGGCTGG 4524  
QY 1021 TGGCGAAGGATAGGCTTGGGAAGGTGTGTTGGACATCTGGGGGGGGGGGG 1080  
Db 4525 TGGCGAAGGATAGGCTTGGGAAGGTGTGTTGGACATCTGGGGGGGGGGGG 4584  
QY 1081 CGCGTGTAGCGTGGCTGGGGGTTACAGTGTGGGGGGGGGGGGGGGGGGGG 1140  
Db 4585 CGCGTGTAGCGTGGCTGGGGGTTACAGTGTGGGGGGGGGGGGGGGGGGGG 4644  
QY 1141 ASGGACGCGCGCACCGGTGACTCAGTCGACTGCTGGGGGGGGGGGGGGGG 1200  
Db 4645 AGGGACGCGCGCACCGGTGACTCAGTCGACTGCTGGGGGGGGGGGGGGGG 4704

QY 1201 AAGAGGCTTACCAAGGTGGAATGGGACTGCTTACGGCTAACCCCTCTGGCTA 1260  
Db 4705 AAAGGCTTACCAAGGTGGAATGGGACTGCTTACGGCTAACCCCTCTGGCTA 4764  
QY 1261 AGGATGTTGGACTGATGCACTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1320  
Db 4765 AGGATGTTGGACTGATGCACTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4824  
QY 1321 AAGCTCTGGGGGATACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1380  
Db 4825 AAGCTCTGGGGGACTACTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4884

QY 1381 GTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1440  
Db 4885 GTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 4944  
QY 1441 CATGTCAAAACGGGTTCACTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1500  
Db 4945 CATGTCAAAACGGGTTCACTGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5004  
QY 1501 GGACATITCCCACATGAGGGATGTCGGGGCTTAACCTGTCAGTGGCAT 1560  
Db 5005 GGAACATITCCCACATGAGGGATGTCGGGGCTTAACCTGTCAGTGGCAT 5064

QY 1561 TATTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1620  
Db 5065 TATTCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5124  
QY 1621 GATTTCACACTAGGACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1680  
Db 5125 GATTTCACACTAGGACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5184  
QY 1681 GCCCCGGGAATCTTCAAGAGGTGGATGGGGGGGGGGGGGGGGGGGGGG 1740  
Db 5185 GCTCCGGAACTCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5244  
QY 1741 AAACCTTCTTCAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1800  
Db 5245 AGGCCTCTTCAAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5304  
QY 1801 TCGGAGCTCCATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1860  
Db 5305 TCACTGGTACATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5364  
QY 1861 CCCTCCACATCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1920  
Db 5365 CCCTCCACATCACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5424  
QY 1921 TGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1980  
Db 5425 TTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5484  
QY 1981 CAAATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2040  
Db 5485 CACCATGCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 5544  
QY 2041 GGC 2043  
Db 5545 GGC 5547

RESULT 10  
US-09-315-850-31  
; Sequence 31, Application US/09315850  
; Patient No. 6217872  
; GENERAL INFORMATION:  
; APPLICANT: OKAYAMA, Hiroto  
; APPLICANT: FUKE, Isao  
; APPLICANT: MORI, Chisato  
; APPLICANT: TAKAMIZAWA, Akahisa  
; APPLICANT: YOSHIDA, Iwao  
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
; TITLE OF INVENTION: DNA AND ANTIGEN POLYPEPTIDE

NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &  
 STREET: 1725 K St. N.W. Suite 1000  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/315,850  
 FILING DATE:  
 APPLICATION NUMBER: US/08/994,686  
 FILING DATE: 01-AUG-1997  
 APPLICATION NUMBER: US 08/224,977  
 FILING DATE: 18-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-210921  
 FILING DATE: 31-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-3105605  
 FILING DATE: 09-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099,706  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/669,996  
 FILING DATE: 02-OCT-1991  
 REFERENCE/DOCKET NUMBER: 900703G  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7917 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..7862  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1..7915  
 OTHER INFORMATION: /note= "sequence = 1500 - 9416 of  
 ; OTHER INFORMATION: SEQ ID NO: 1"  
 US-09-315-850-31

Query Match Best Local Similarity 87.9%; Score 1796.6; DB 3; Length 7917;  
 Matches 1889; Conservative 92.5%; Prev. No. 0; Mismatches 154; Indels 0; Gaps 0;

Qy 1 TGGAGCCGGCTTCTCACAGGCCCTACCCACGGTGGATGCCACTTCCTCGTCCAAACRAAG 60  
 Db 3505 TGGAGAGGTGTCACAGGCCCTACCCATATAGTCACACTTCTGTCCCAGACCAAG 3564  
 Qy 61 CAGGACGCCGGCAGCAACTTCCCTAACCTCGTGCAGCTACAGCTACCCAACTGTTG 120

---

Db 3365 CAAGCAGAAGACACTTCCCTACTGTAGCATACCAAGCCACGGGTGGCCAGGGCT 3624  
 Qy 121 CAGGCCACCTCCATATGGATCAAATGGGAAGGTCTCATACGCCATAAGCTACT 180  
 Db 3625 CAGGCCACCTCCATATGGATCAAATGGGAAGGTCTCATACGCCATAAGCTAC 3684  
 Qy 181 CTGGCGCCGGCCACACCCCTGCCTGATAGGCTGGACGCCCTGGCTCAAGCTAC 240  
 Db 3685 CTGGCGCCGGCCACACCCCTGCCTGATAGGCTGGACGCCCTGGCTCAAGCTAC 3744  
 Qy 241 ACACACCCATACCAATTCTCATGCATCAGCTGGCTGGAGGGCTGAC 300  
 Db 3745 ACCACACCCATACCAATTCTCATGCATCAGCTGGCTGGAGGGCTGAC 3804  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/994,686  
 FILING DATE: 01-AUG-1997  
 APPLICATION NUMBER: US 08/224,977  
 FILING DATE: 18-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-210921  
 FILING DATE: 31-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-3105605  
 FILING DATE: 09-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099,706  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/669,996  
 FILING DATE: 02-OCT-1991  
 REFERENCE/DOCKET NUMBER: 900703G  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7917 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..7862  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1..7915  
 OTHER INFORMATION: /note= "sequence = 1500 - 9416 of  
 ; OTHER INFORMATION: SEQ ID NO: 1"  
 US-09-315-850-31

Db 301 AGCACCTGGCTGGGGGCTCTGAGCTGGCTGGCTGGGATTTGTGAC 360  
 Db 3105 AGCACCTGGCTGGGGGCTCTGAGCTGGCTGGGATTTGTGAC 3864  
 Qy 421 AGCGAACCTCTTACCGGGAGTCTGATGAGATGAAAGAGTCGGCTTGAC 480  
 Db 361 GGCAGCTGGTCATTGGGGTAGATCTGTGAGCTGGCTGGGGCTATGTTCCGC 420  
 Db 3865 GGCAGCTGGTCATTGGGGTAGATCTGTGAGCTGGCTGGGGCTATGTTCCGC 3924  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-210921  
 FILING DATE: 31-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-3105605  
 FILING DATE: 09-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099,706  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/669,996  
 FILING DATE: 02-OCT-1991  
 REFERENCE/DOCKET NUMBER: 900703G  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 31:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7917 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..7862  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: 1..7915  
 OTHER INFORMATION: /note= "sequence = 1500 - 9416 of  
 ; OTHER INFORMATION: SEQ ID NO: 1"  
 US-09-315-850-31

Db 4105 GAGACATCTGGGGAGACATGTCGACTTCATGGGGATACGGCTGGAC 660  
 Db 4145 ACAGCCACCAACAGGGAGGCTGCTCCGAGCAATCAAGGAAAGGCCTCGGCT 600  
 Db 4165 TTATCCCTCTGCCTGGAGATCCGCATGATCACTGATGGGGATGGCTCCAGTGGAGCC 4104  
 Qy 601 GAGACCTCTGGGGAGACATGTCGACTTCATGGGGATACGGCTGGAC 4044  
 Db 4165 TTATCCCTCTGCCTGGAGATCCGCATGATCACTGATGGGGATGGCTCCAGTGGAGCC 4224  
 Qy 721 ACTAGCCGCTTACCAACCCAACTACCCCTCCGCTTAACATCTGGGGATGGCT 720  
 Db 4225 ACCGCGCGCTTACCAACCCAACTACCCCTCCGCTTACATCTGGGGATGGCT 4284  
 Qy 781 GCACAACCTGCTCCCGCAAGTGTGCTCACTGGGGATGGCT 840  
 Db 4285 GCCAACCTCGCCCCCGACGCCGCTGGGGATGGCT 4344  
 Qy 841 GCTGTTGCCAGCATAGGCCCTGGAGGTCTGTCAGACATCTGGGGCTATGGCA 900  
 Db 4345 GCTGTTGCCAGCATAGGCCCTGGAGGTCTGTCAGACATCTGGGGCTATGGCA 4404  
 Qy 901 GAGGTGCGCAGGGGGCTTGGTGGCTTAAGGTCTGAGGCCGAATGGCCCTCACGGAG 960  
 Db 4405 GGAGTGGCGGGGGCTTGGTGGCTTAAGGTCTGAGGCCGAATGGCCCTCACGGAG 4464  
 Qy 961 GACCTGCTTAACCTACTCCCTGCCATCTCTCCCTGGCTGGGGCTGGCT 1020  
 Db 4465 GACCTGCTTAACCTACTCCCTGCCATCTCTCCCTGGCTGGGGCTGGCT 4524  
 Qy 1021 TCGGACGGATCTGGCGACGGTGGCTGGAGGGGGCTGGCTGGATTAAC 1080  
 Db 4525 TGTGACCAACTCTGGCTGGCTGGAGGGGGCTGGCTGGATTAAC 4584  
 Qy 1081 CGCTGTAGCTTCGCTCGGGGGTAACTATGTTCCGCCACGACTATGTGCCAGAG 1140  
 Db 4585 CGCTGTAGCTTCGCTCGGGGGTAACTATGTTCCGCCACGACTATGTGCCAGAG 4644  
 Qy 1141 AGGACGCCGGCAGCAACTCTGGCTGGAGGGGGCTGGCTGGATTAAC 1200



QY	121	CAGGCCACACTCCATCATGGCATCATGGGAACTGGGAAGTGTCTACCGCTTAAGCTACT	180	QY	1201	AAGAGSTCCACCACTGGATTACGGGACTCTCCAGCCCTGGCTCGGCTCGGCTCGGCTA	1260
Db	4792	CTGGCCACACTCCATCATGGCATCATGGGAACTGGGAAGTGTCTACCGCTTAAGCTACT	4851	Db	5932	AAAAGSCTCCACCACTGGATTACGGGACTCTCCAGCCCTGGCTCGGCTCGGCTCGGCTA	5931
QY	181	CTGGCGGGCACACCCCTCTGTATAGGTGGGGGGCTGTCAAAACGAGCTCACCTC	240	QY	1261	AGGATGTTGGGACTGGATATGCAGGTTGGGACCTGACTTCAGAGCTGCTGGCTCGGCT	1320
Db	4852	CTGCAAGGGCACACCCCTCTGTATAGGTGGGGGGCTGTCAAAACGAGCTCACCTC	4911	Db	5932	AGGATGTTGGGACTGGATATGCAGGTTGGGACCTGACTTCAGAGCTGCTGGCTCGGCT	5991
QY	241	ACACACCCATACCAATTATCATGGCATCATGGGACTCTGGAGCTGGCTACGAGTCAC	300	QY	1321	AAGCTCTGGGGGATACCGGASTCCCTTTCATGCCAACGGTACACTTCAGAGCTGCTGG	1380
Db	4912	ACCCACCCATACCAATTATCATGGCATCATGGGACTCTGGAGCTGGCTACGAGTCAC	4971	Db	5992	AAGCTCTGGGGGATACCTGGAGCTGGCTACGAGTCACGAGTCACGAGTCACGAGTC	6051
QY	301	AGCACCTGGGCTGGGGGGCTGTCAGGTTGGGACCTGACTTCAGAGCTGCTGG	360	QY	1381	GCTGGGGGGACGGCATGCGAGGACCTGCTCGGAGGACATGGCTCGGCTACGGGA	1440
Db	4972	AGCACCTGGGCTGGGGGGCTGTCAGGTTGGGACCTGACTTCAGAGCTGCTGG	5031	Db	6052	GCTGGGGGGACGGCATGCGAGGACCTGCTCGGAGGACATGGCTCGGCTACCGGA	6111
QY	361	GCGAGCTGGTCTGGGATGATCATCTGTGCGGGGCGGCGCTATGTTCCCGAC	420	QY	1441	CATGTCAAAACGGTTCATGAGGATGAGTGGCTGGAGAGTGCTGGAGCTCCCTAC	1500
Db	5032	GCGAGCTGGTCTGGGATGATCATCTGTGCGGGGCGGCGCTATGTTCCCGAC	5091	Db	6112	CATGTCAAAACGGTTCATGAGGATGAGTGGCTGGAGAGTGCTGGAGCTCCCTAC	6171
QY	421	AGGGAGTCTCTACCGGAACTGAGTAGGATATCTGTGCGGAGCTCCCTAC	480	QY	1501	GAACATTCGGGAGCTGGGAGCTGGGAGGAGTGGCTAGAAGCTGTTAGTAAC	1560
Db	5092	AGGGAGTCTCTACCGGAACTGAGTAGGATATCTGTGCGGAGCTCCCTAC	5151	Db	6172	GAACATTCGGGAGCTGGGAGGAGTGGCTAGAAGCTGTTAGTAACATGTCAC	6231
QY	481	ATCGAGGAGGAAATGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	540	QY	1561	TATTCGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1620
Db	5152	ATCGAGGAGGAAATGAGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5211	Db	6232	TATTCGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	6291
QY	541	ACAGCCACCAAGCAACGGGAGGGGGGGGGGGGGGGGGGGGGGGGGGG	600	QY	1621	GATTCACTAGTGAGGAGATGACCACTGACAACGTAAGTAATGCGGTGCA	1680
Db	5212	ACAGCCACCAAGCAACGGGAGGGGGGGGGGGGGGGGGGGGGGGGG	5271	Db	6292	GATTCACTAGTGAGGAGATGACCACTGACAACGTAAGTAATGCGGTGCA	6351
QY	601	GAGACCTCTGGGAAACACATGAGACTCATGAGGGATACTGAGTGGAGC	660	QY	1681	GCCCCGGAATCTTACAGAAGTGGATGGGGGGGGGGGGGGGGGGGG	1740
Db	5272	GAGACATTCGGGAGGAAATGAGCTGGGGGGGGGGGGGGGGGGGG	5331	Db	6352	GCTCCCTAACCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	6411
QY	661	TGTCCACTCTGGGAATCCGGGATGATCATGATGGGTCAAGCCCTGTC	720	QY	1741	AAACCTCTTACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1800
Db	5332	TATCCACTCTGGGAACCCGAATACATCATGATGGCATTCAGCCCTAC	5391	Db	6412	AGGCCTCTTACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	6471
QY	721	ACTAGCCGCTACACCCAAATCTACCTCTGTGTCACACCTCTGGGGGG	780	QY	1801	TCGAGGCTCATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1860
Db	5392	ACGAGCCGCTACACCCAAAGTACCTCTGTGTCACACCTCTGGGGGG	5451	Db	6472	TCACACCTACATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	6531
QY	781	GCCCCACTCTCCCGGAGCTCTACCTCTGAGGGGGGGGGGGGGGG	840	QY	1861	CCCTCCCACATCACAGGAGAACGGGAGTAGCTGGCTGGCTGG	1920
Db	5452	CCCAAATCTCGGCCCCAACGGGCGCTCGGCCCCAACGGGCGCTCG	5511	Db	6532	CCCTCCCACATCACAGGAGAACGGGAGTAGCTGGCTGGCTGG	6591
QY	841	GCTGTGGCACCATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	900	QY	1921	TTGGCGGCTCTTACCTAGGCTGGGGGGGGGGGGGGGGGGGGGG	1980
Db	5512	GCTGTGGCACCATGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5571	Db	6592	TTGGCGGCTCTTACCTAGGCTGGGGGGGGGGGGGGGGGGGG	6651
QY	901	GGAGTGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	960	QY	1981	CAAAATGACTCTCCAGGGCTGGGGGGGGGGGGGGGGGGGGGG	2040
Db	5572	GGAGTGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	5631	Db	6652	CACCATCTCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	6711
QY	961	GACCTGTTACTTACTTCCCTGCCATCTCCCTCTGGGGGGGGGG	1020	QY	2041	GCC 2043	
Db	5632	GACCTGTTACTTACTTCCCTGCCATCTCCCTCTGGGGGGGGGG	5691	Db	6712	GCC 6714	
QY	1021	TCCGCGGGGTTACTGGCTGGGGGGGGGGGGGGGGGGGGGGGG	1080	RESULT 12			
Db	5692	TGTGAGGAAATACTCTGGGCACTGGGGGGGGGGGGGGGGGG	5751	US-08-384-616-13			
QY	1081	GGGCTGATAGGTTCTCCCTGGGGGGGGGGGGGGGGGGGGGG	1140	Sequence 13, Application US/08384616			
Db	5752	GGGCTGATAGGTTCTCCCTGGGGGGGGGGGGGGGGGGGGGG	5811	; Patent No. 5,647,01			
QY	1141	AGCGAGCCGAGGAGCTCTCTGGGGGGGGGGGGGGGGGGGG	1200	; GENERAL INFORMATION:			
Db	5812	AGCGAGCCGAGGAGCTCTCTGGGGGGGGGGGGGGGGGGGG	5871	; APPLICANT: OKAYAMA, Hiroto			
			; APPLICANT: FUKE, Iwao				
			; APPLICANT: MORI, Chisato				
			; APPLICANT: TAKAMIZAWA, Akahisa				
			; APPLICANT: YOSHIDA, Iwao				
			; TITLE OF INVENTION: NON-B HEPATITIS VIRUS GENOMIC				

NUMBER OF SEQUENCES: 50  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Naughton  
 STREET: 1725 K St. N.W. Suite 1000  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
 COMPUTER: IBM PC compatible  
 CURRENT APPLICATION DATA:  
 FILING DATE:  
 CLASSIFICATION: 424  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/1769, 996  
 FILING DATE: 02-OCT-1991  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 31-AUG-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stevens-Smith, Thereea M.  
 REGISTRATION NUMBER: 36,284  
 REFERENCE/DOCKET NUMBER: 900703B  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 TELEX: 440142

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 9030 base pairs  
 TYPE: nucleic acid  
 STRANDBNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA

FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..9030  
 OTHER INFORMATION: /note: "sequence = 333 - 9362 of  
 OTHER INFORMATION: SEQ ID NO: 1"

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..9030  
 ; US-08-616-13

Query Match 87.9%; Score 1796.6; DB 2; Length 9030;  
 Best Local Similarity 92.5%; Pred. No: 0;  
 Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 1 TGGGACGGCGCTTCAAGGCTCACCCACCTGGATGCCACTCTGTCCAAACAGG 60  
 Qy 61 CAGGCAGGAGAACCTCCCTACTGGATAATGGAAGTGTCTATACGGCTAAAGCC 120  
 Db 4732 CAGGCCCACTCCATATGGATAATGGAAGTGTCTATACGGCTAAAGCC 4791  
 Qy 121 CAGGCCCACTCCATATGGATAATGGAAGTGTCTATACGGCTAAAGCC 180  
 Db 4792 CAGGCCCACTCCATATGGATAATGGAAGTGTCTATACGGCTAAAGCC 4851

Qy 181 CTGGCGGCGCAACACCCCTGGTGTATAGCTGGAGCGCTCCAAACAGGACCC 240  
 Db 4852 CTGACGGGCCAACACCCCTGGTGTACAGCTGGCGCTGGAGCGCTCCAAACAGGACCC 4911  
 Db 241 ACACACCCATAACCAATTACATGTGGCATGTGAGTCAGCCGACCTGGAGTCAG 300  
 Qy 4912 ACCAACCCATAACCAATTACATGTGGCATGTGAGTCAGCCGACCTGGAGTCAG 4971  
 Qy 301 AGCACCTGGGTGCTGGGGGGCTCTGGCAGCTCTGGCTGTGCTGAGTCAG 360  
 Db 4972 AGCACCTGGGTGCTGGGGGGCTCTGGCAGCTCTGGCTGTGAGTCAG 5031  
 Qy 361 GGCACTGGCATGGTAGATCATGTGGCTGAGTCAGCCGACCTGGAGTCAG 420  
 Db 5032 GGCACTGGGTGCTGGTAGATCATGTGGCTGAGTCAGCCGACCTGGAGTCAG 5091  
 Qy 421 AGGGAGTCTCTACAGAGGTGATGAGATGGAGAGTCAGCCGACCTGGAGTCAG 480  
 Db 5092 AGGGAGTCTCTACAGAGGTGATGAGATGGAGAGTCAGCCGACCTGGAGTCAG 5151  
 Qy 481 ATCGAAGCAGGATGCGCTGGCCAGACTGCGAGCTTCAGCAAAGGCTCGGTCTGGAG 540  
 Db 5152 ATCGAAGCAGGATGCGCTGGCCAGACTGCGAGCTTCAGCAAAGGCTCGGTCTGGAG 5211  
 Qy 601 GAGACCTCTGGCAACACATGTTGAACTTCATGGGGATACAGTACTAGAAC 660  
 Db 5212 GAGACATCTGGCAAAGGAGGCTCTGTCGGGGAGTCAGAAGGCTCGGTCTGGAG 5331  
 Qy 541 ACAGGCCAACACAGGGAGGCTCTGTCGGGGAGTCAGAAGGCTCGGTCTGGAG 600  
 Db 5212 ACAGGCCAACACAGGGAGGCTCTGTCGGGGAGTCAGAAGGCTCGGTCTGGAG 5271  
 Qy 661 TTGTCCACTCTGGCTGGATCCCGGAGTGACACTGATGGGTTCAGGCTCTGG 720  
 Db 5332 TTATCCATCTGGCTGGAAACCCGGCAATAGCATCTGATGGCATTCAGGCT 5391  
 Qy 721 ACTAGCCGCTACACCACTACCCCTCTGCTGTTAACCTCTGGGGATGGTAGCC 780  
 Db 5392 ACCACCCGCTCACCCACCAAGTACCTCTGGTAACTCTGGGGATGGTAGCC 5451  
 Qy 781 GCCCAACTCTCCCCAACCTGCTGCTTAAGTCTGAGCTTGTAGGGCGGCGATGTCGGTGG 840  
 Db 5452 GCCCAACTCTCCCCAACCTGCTGCTTAAGTCTGAGCTTGTAGGGCGGCGATGTCGGTGG 5511  
 Qy 841 GCTGTGCGACGATGGCTTGGAAGGCTCTGGAGACATCTGGGGCTATGGCA 900  
 Db 5512 GCTGTGCGACGATGGCTTGGAAGGCTCTGGAGACATCTGGGGCTATGGCA 5571  
 Qy 901 GGAGTGGCGAGCGCGCTGGGGCTTAAAGTCATGGGGGAATGCCCTCCAGG 960  
 Db 5572 GGAGTGGCGAGCGCGCTGGGGCTTAAAGTCATGGGGGAATGCCCTCCAGG 5631  
 Qy 961 GACCTGGTTAACTTACTCCCTGGCCATCTCTCTGGCTGGGGCTGGTGGGGTGTG 1020  
 Db 5632 GACCTGGTTAACTTACTCCCTGGCCATCTCTCTGGGGCTGGTGGGGTGTG 5691  
 Qy 1021 TCCGCAGGATACTGGCACCTGGTCAAGGGAGGGGGCTGGCATGTGAA 1080  
 Db 5692 TGTGAGCAGTACTGGCTGGGGGAAATCATGTTCCCACGGCTATGTGAA 5751  
 Qy 1081 CGGTGTAGGGTGGCTGGGGGATACATGTTCCCACGGCTATGTGAA 1140  
 Db 5752 CGGTGTAGGGTGGCTGGGGGAAATCATGTTCCCACGGCTATGTGAA 5811  
 Qy 1141 AGCGACGCCAGCACGCTGGTCACTGATGTTCTGGGGCTGGGGCTGGTGTG 1200  
 Db 5812 AGCGACGCCAGCACGCTGGTCACTGATGTTCTGGGGCTGGGGCTGGTGTG 5871  
 Qy 1201 AGAGCTCCACCGAGGATTAAGGGAGTGTCTGCTCAAGGCTGCTGGGGCTGA 1260  
 Db 5872 AAAAGCTCCACCGAGGATTAATGAGCTGCTGCTGCTGCTGCTGCTGCTG 5931  
 Qy 1261 AGGAGTGGTGGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320

CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 in, 1.44MB  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/904,686A  
 FILING DATE: 01-AUG-1997  
 APPLICATION NUMBER: US 08/324,977  
 FILING DATE: 18-OCT-1994  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 31-AUG-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-305605  
 FILING DATE: 09-NOV-1990  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099,706  
 FILING DATE: 30-JUL-1993  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/769,996  
 FILING DATE: 02-OCT-1991  
 PRIORITY APPLICATION DATA:  
 APPLICATION NUMBER: US 07/635,451  
 FILING DATE: 28-DEC-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: McLeland, Le-Nhung  
 REGISTRATION NUMBER: 31,541  
 REFERENCE/DOCKET NUMBER: 900703G  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202) 659-2930  
 TELEFAX: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9030 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: m<sub>1</sub>sc feature  
 LOCATION: 1..9030  
 OTHER INFORMATION: /note: "sequence = 333 - 9362 of  
 OTHER INFORMATION: SEQ ID NO: 1"  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..9030  
 ;US-08-904-686A-13

Query Match 87.9%; score 1795 6; DB 2; Length 9030;  
 Best Local Similarity 92.5%; Pred. No. 0; Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0  
 ; GENERAL INFORMATION:  
 ; Sequence 13, Application US/08904686A  
 ; Patent No. 598130  
 ; APPLICANT: OKAYAMA, Hiroto  
 ; FUKU, Iiao  
 ; APPLICANT: MORI, Chisato  
 ; APPLICANT: TAKAMIZWA, Akahisa  
 ; APPLICANT: YOSHIDA, Iwao  
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
 ; NUMBER OF SEQUENCES: 50  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSE: Armstrong, Westerman, Hattori, McLeland &  
 ; STREET: 1725 K St. N.W. Suite 1000

QY 181 CTGCGCGGGGAAACGCCCTTCTGTATAGSTGGAGGCCGCCAACAGGGTCACCCCTC 240  
Db 4852 CTGCACGGCCAACACCCCTCTGTACAGCTGGAGCCGCTCAGATGAGTCACCTC 4911  
QY 241 ACACACCCCATAACCAATTATCATGGCATCATCGCGACCTGGGTCAGTCAG 300  
Db 4912 ACCACACCCATAACCAATTATCATGGCATCATCGCGACCTGGGTCAGTCAG 4971  
QY 301 AGCACCTGGGTTCTGGGGGTTCTTCAGCTCTGGGTGCGTATTGGTACA 360  
Db 4972 AGCACCTGGGTTCTGGGGGAGTCCTCTGAGCTCTGGCGGTATGGTACA 5031  
QY 361 GCGCAGCGTGTGTTCTGGGAGGTCTGGTCA 420  
Db 5032 GCGCAGTGTTGTTCTGGGAGGTCTGGTCA 5091  
QY 421 AGGAAGTCTCTTACAGGAGTCTGGTCA 480  
Db 5092 AGGGAGACTCTTACAGGAGTCTGGTCA 5151  
QY 481 ATCAGCAGGGAAATGAGCTGGCAGCTGCCAGTCAGTCAGCAA 540  
Db 5152 ATCAGCAGGGAAATGAGCTGGCAGTCAGCAA 5211  
QY 541 ACAGCCACCAAGCAAGCGGAGGCCCTGCGCCGCGTGGAGTCAGTGAGCCT 600  
Db 5212 ACAGCCACCAAAACAGGGAGCTGCGCCGAGCAATCAAGCAGAAGCGCTCGGGTACTGCAA 5271  
QY 601 GAGACCTTCTGGGAAACAGATGTGACTCATAGCGGAGTACAGTAGCTAGGAGC 660  
Db 5272 GAGACATTCTGGGAAAGCAATGAGTCA 5331  
QY 661 TGTGCACTCTGCTGGAAATCTCCGGATTCATCATCTGATGGCGTCAAGCCCTGTC 720  
Db 5332 TTATCCACTCTCTGGAAACCCGGAAATAGCATCATGATGGCATTCACGCCCTAC 5391  
QY 721 ACTAGCCGCGCTCACCCCAAATCTACCTCTCTGGTACACCTGGGATGGTAGCC 780  
Db 5392 ACCAGCCGCGCTCACCCCAAATGACTCCCTCTGGTACACCTGGGATGGTAGCT 5451  
QY 781 GCCCAACTCTGCTCCCGAGCTGCTTCAGTTCTGTTAGCCGGCATCTGGTGG 840  
Db 5452 GCCCAACTCTGCCCTCCAGGCCCTTCAGCTGGGCGCATGCCGGTGG 5511  
QY 841 GCTGTGCGACGATAGCCCTTGGAAAGGTGTTGAGCATCTGGGGGATGGAGCA 900  
Db 5512 CCTGTGCGACGATAGCCCTTGGAAAGGTGTTGAGCATCTGGGGGATGGAGCA 5571  
QY 901 GGAGTGCGAGCGGCGCTGGCCTTAAGGTGAGCTGAGCGCGGAATGCCCTCACCAG 960  
Db 5572 GGAGTGCGCGCGCTGGCCTTAAGGTGAGCTGAGCGCGGAATGCCCTCACCAG 5631  
QY 961 GACCTGGTTAACTTACTCTCCATCTCTCTGGTGCCTGTCGGGGTCTG 1020  
Db 5632 GACCTGGTCAATCCTCTCCATCTCTCTGGTGCCTGTCGGGGTCTG 5691  
QY 1021 TCGCAGCGATAGCTGCTCGGAGCTGGTCAAGGGAGGGGGCTGCACTGGATGAC 1080  
Db 5692 TGTGCACTAACATGCTGCTGAGCTGGTCAAGGGAGGGGGCTGCACTGGATGAC 5751  
QY 1081 CGGCTGTAGTGTCTGCTCGGGGTAACCATGTTCCCAAGCAGTGTGCGAGAG 1140  
Db 5752 CGGCTGTAGTGTCTGCTCGGGGTAATGTTGCTGCGCTGCGCTGCGCTGCGCTG 5811  
QY 1141 AGCGAGCGCGAGCAGCTGTACTCAGATCTCGACCTACTATCACCACAGTGTG 1200  
Db 5812 AGCGAGCGCGAGCAGCTGTACTCAGATCTCGACCTACTATCACCACAGTGTG 5871  
QY 1201 AGAGGAGTCCACCCAGTGTGATTACGGAGCTCTCCAGCGCTGCTCCGGCTGCGCTA 1260  
Db 5872 AAAGAGTCCACCCAGTGTGATTACGGAGCTCTCCACCGCTGCTCCGGCTGCGCTA 5931  
QY 1261 AGGGAGTGTGTTGGACTGTGATATGCCAGCTTGTGACTCTCAAGACCTGGCTCCAGTCC 1320

---

RESULT 14  
US-09-315-850-13  
; Sequence 13, Application US/09315850  
; Patent No. 6,217,872  
; GENERAL INFORMATION:  
; APPLICANT: OKAYAMA, Hiroto  
; APPLICANT: FUKE, Isao  
; APPLICANT: MORI, Chisato  
; APPLICANT: TAKAMIZAWA, Akahisa  
; APPLICANT: YOSHIDA, Iwao  
; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
; TITLE OF INVENTION: DNA AND ANTIGEN POLYPEPTIDE  
; NUMBER OF SEQUENCES: 50  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &  
; STREET: 1725 K St. N.W. Suite 1000

CITY: Washington  
 STATE: D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20006  
 COMPUTER READABLE FORM:  
 OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/315, 850  
 FILING DATE:  
 APPLICATION NUMBER: DATA:  
 APPLICATION NUMBER: US/08/904, 686  
 FILING DATE: 01-AUG-1997  
 APPLICATION NUMBER: US 08/324, 977  
 FILING DATE: 18-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-167466  
 FILING DATE: 25-JUN-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-230921  
 FILING DATE: 31-AUG-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 2-305605  
 FILING DATE: 09-NOV-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/099, 706  
 FILING DATE: 30-JUL-1993  
 PRIOR APPLICATION DATA:  
 NAME: McLeland, Le-Nhung  
 REGISTRATION NUMBER: 31, 541  
 REFERENCE/DOCKET NUMBER: 9000703G  
 TELECOMMUNICATION INFORMATION:  
 ATTORNEY/AGENT INFORMATION:  
 TELEPHONE: (202) 887-0357  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 9030 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA from genomic RNA  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 1..9030  
 OTHER INFORMATION: /note: "sequence = 333 - 9362 of"  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 1..9030  
 US-09-315-850-13

Query Match 87.9%: Score 1796; DB 3; Length 9030;  
 Best Local Similarity 92.5%; Pred. No. 0; Mismatches 154; Indels 0; Gaps 0;  
 Matches 1889; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

Qy 1 TGGGAGGAGGAACTCCCTACTGTGGTACAGGCTACCCACCTGGATCAGGCCACTTCCCTGTCACAAACAAAG 60  
 Db 4672 TGGGAGGAGGAACTCCCTACTGTGGTACAGGCCACTTCCCTGTCACAAACAAAG 4731

Qy 61 CAGGAGGAGGAACTCCCTACTGTGGTACAGGCCACTTCCCTGTCACAAACAAAG 120  
 Db 4732 CAGGAGGAGGAACTCCCTACTGTGGTACAGGCCACTTCCCTGTCACAAACAAAG 4791

Qy 121 CAGGCCCACTCCATCATGGGATCAATGTTGAGGTCTCATACCGCTTAAGGCTACT 180  
 Db 4972 AGGAACTGGTCTTACGGGAGGTGAGNTGAGNTGAGAGTGGAGTGGCGGACTCGCTGAGCTGGAGCTGGAC 5031

Db 4972 AGGAACTGGTCTTACGGGAGGTGAGNTGAGNTGAGAGTGGAGTGGCGGACTCGCTGAGCTGGAGCTGGAC 4911

Db 241 AACACACCCATACCAAATTCTATCGCATGCATCATGCTAGCGGACCTGGAGCTGGAC 300

Db 361 GGCAGCCCTGGTCTGTCATGTCGGTAGATCTTGTCGGCGGCGCGGATATGTTCCGAC 420

Db 5032 GGAGTGGTCTGTCATGTCGGTAGATTCTGTCCGGAGCCGGACATTGTCGGAC 5091

Qy 481 ATGAGGAGGGATGCGCTCGCGAGCACTGAGAATGGAAAGGGCTGGGTGCTCAG 540

Db 5152 ATGAGGAGGGATGCGCTCGCGAGCACTGAGAATGGAAAGGGCTGGGTGCTCAG 5211

Db 541 ACGGCCAACAAAGGGGAGGCTGGAGTCCAGTGGAGCTGATGCGGACTCCCTAC 480

Db 5272 GAGACATTCCTGGCGGAACATGTCATGTCGAATTCTAGGGAGCTGGACCTCCCTAC 5151

Db 661 TTGTCACCTGGCTGGAAATCCGGCAATGATCACATGATGGGTACAGCTTGTC 720

Db 5332 TTATCCCTCTGGAAACCCGGATAGCATGTCATGTCATGTCATGTCATGTC 5391

Db 721 ACTAGCCGCTCACCACTACCCCTCTGCTAACATCCTGGGGATGGGTGCC 780

Db 5392 ACCAGCCGCTCACCACTACCCCTCTGCTAACATCCTGGGGATGGGTGCC 5331

Qy 781 GCCAACATGCCCTCCAGCTGCTGTCAGTTGCTAGGGCGCATGTCATGTC 840

Db 5452 GCCTAACATGCCCTCCAGCCCGCTCTGCTAACATCCTGGGGATGGGTGCC 5511

Qy 841 GCTGTTGCAAGATAGGCTTGGAAAGGGCTGTGAGACATCTTGGGGCTATGGCA 900

Db 5512 GTGTTGCAAGATAGGCTTGGAAAGGGCTGTGAGACATCTTGGGGCTATGGCA 5571

Qy 901 GGAGTGGCAGGCGCTGCTGGCTGGCTTAAGGTGATGGCGGAAATGGCTTGTGGAG 960

Db 5572 GGAGTGGCAGGCGCTGCTGGCTGGCTTAAGGTGATGGCGGAGATGGCTTGTGGAG 5631

Qy 961 GACCTGTTAATCTACCTCCCTCCATCTCTCTCCCTGGCTGCTGCTGGGTGTGCTG 1020

Db 5632 GACCTGTTAATCTACCTCTCCCTGGCTGCTGGGTGTGCTGCTGGGTGTG 5691

Qy 1021 TGGGAGGAGTACTGCTGGGAGCTGGTCAAGGGAGGGGGCTGTGAGGTGAC 1080

Db 5592 TGGGAGGAGTACTGCTGGGAGCTGGTCAAGGGAGGGGGCTGTGAGGTGAC 5751

Qy 1081 CGCTGTTAGCTTCGCTCAGGGGAAACCTGTTCCACGCACTATGTCGAG 1140

Db 5752 CGCTGTTAGCTTCGCTCAGGGGAAACCTGTTCCACGCACTATGTCGAG 5811

Qy 1141 ACGGACCCGGAGGAGCTGGTCAAGGGAGGGGGCTGTGAGGTGAC 1200

Db 5812 AGGACCCGGAGGAGCTGGTCAAGGGAGGGGGCTGTGAGGTGAC 5871

Qy 1201 AAGAGGCTTACGGAGGAGTACAGGAGGAGCTCCACGGCCCTGGCTGGCTA 1260

Db 5872 AAAAGGCTTACGGAGGAGTACAGGAGGAGCTCCACGGCCCTGGCTGGCTA 5931

ADDRESSEE: Naughton  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/324,977  
FILING DATE: 18-OCT-1994  
PRIORITY NUMBER: JP 2-230921  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIORITY NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIORITY NUMBER: US 08/099,706  
FILING DATE: 30-JUL-1993  
PRIORITY NUMBER: US 07/769,996  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens-Smith, Theresa M.  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: 900703D

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
TELEX: 440142

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9416 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA from genomic RNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 333..9362

US-08-324-977-1

Query Match 87.9%; Score 1796.6; DB 1; Length 9416;  
Best Local Similarity 92.5%; Pred. No. 0; Mismatches 154; Indels 0; Gaps 0;  
Matches 1889; Conservative 0;

QY 1 TCGGAGGGCTTCAGGGCTCACCCACGTGGTGCCTCTGTCCAAACAAG 60  
Db 5004 TGGGAGATGTTCTTACAGGCTTACCCACCTTGCTGGAGTCAAG 5063  
QY 61 CAGGAGGGAGAACCTTACCTGGTGGAGTACAGGTACTGTGGCTAGGCC 120  
Db 5064 CAGGAGGGAGAACCTTACCTGGTGGAGTACAGGTACTGTGGCTAGGCC 5123  
QY 121 CAGGCCAACCTCATGATGGTGAATGGGAATGGTGTCTCATGGTAAAGCT 180  
Db 5124 CAGGCCAACCTCATGATGGTGAATGGGAATGGTGTCTCATGGTAAAGCT 5183  
QY 181 CTGGCGGGCAACACCCCTGTTAGCTGGGAGCCGCTCAAACAGGGTCACCTC 240  
Db 5184 CTGCAAGGGCAACACCCCTGTTAGCTGGGAGCCGCTCAAACAGGGTCACCTC 5243  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &

RESULT 15  
US-08-324-977-1  
Sequence 1, Application US/08324977  
; Parent No. 574739  
; GENERAL INFORMATION:  
; APPLICANT: OKAYAMA, Hiroto  
; APPLICANT: FUKE, Isao  
; APPLICANT: MORI, Chisato  
; APPLICANT: TAKAMIZAWA, Akahisa  
; APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: cDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, McLeland &



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

### OM nucleic - nucleic search, using sw model

Run on: February 25, 2005, 03:07:49 ; Search time 6005 Seconds

12950.092 Million cell updates/sec

Title: US-09-664-363-20

Perfect score: 2043

Sequence: 1 TGGAGGGGCTTCACAGG.....TGTGGCCGCATGAGATGGC 2043

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:\*

1: gb\_elt1:\*

2: gb\_elt2:\*

3: gb\_hac:\*

4: gb\_elt3:\*

5: gb\_elt4:\*

6: gb\_elt5:\*

7: gb\_elt6:\*

8: gb\_gb81:\*

9: gb\_gb82:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No. Score Query Length DB ID Description

1 49.2 2.4 925 9 CNS0091P

AL053013 Drosophila

LOCUS CNS0091P 925 bp DNA linear GSS 03-JUN-1999

DEFINITION Drosophila melanogaster genome survey sequence TFR3 end of BAC #

fly, genomic survey sequence.

ACCESSION AL053013.1

VERSION GI:4934461

KEYWORDS GSS.

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydriidae; Drosophilidae; Drosophila.

REFERENCE 1 \{bases 1 to 925\}

AUTHORS Genoscope

TITLE Direct Submission

JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

Web : [www.genoscope.cns.fr/](http://www.genoscope.cns.fr/)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see <http://www.fruitfly.org>. The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

FEATURES source

organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/db\_xref="taxon:7227"

/clone="BACR9D16"

/clone\_id="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match

Best Local Similarity

13.9%;

Pred. No. 0.17;

Length 925;



RESULT 5

AUTHORS Cruz, A.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (17-JUN-2003) Cruz A.K., University of Sao Paulo,  
 Department of Molecular and Cell Biology, FMRP, Avenida  
 Bandeirantes 3900, Ribeirao Preto, SP, 14049-900, BRAZIL  
 COMMENT Clone requests: akcruz@mrp.usp.br.

FEATURES source

1. .635  
 /organism="Leishmania braziliensis"  
 /mol\_type="genomic DNA"  
 /strain="MWHO/BR/75/M2904"  
 /db\_xref="taxon:5600"  
 /clone="LBAF018C06"

ORIGIN

Query Match 2.3%; Score 47; DB 9; Length 635;  
 Best Local Similarity 48.7%; Pred. No. 0.61; Mismatches 0; Indels 0; Gaps 0;  
 Matches 128; Conservative 0; MisMatches 135; Indels 0; Gaps 0;

Qy 780 CGCCCAACTCTCCCGGAGTGCCTTCAGCTTGAGGGGGCATATGGAGTC 839  
 /CGCGCACTCCACGTGCCCTGCTGCTTGAGCCAGGGTCCACTGCCACGCTGC 323

Qy 840 GGCCTGTGGAGCATAGGCCCTGGAGGTCCTGGGACATGGGGCATGGAGC 899  
 /GGGAGGGGAGGCGGCTCTGGGCTTAAGTCAGTGGAGCTGGAGCTGGGAGA 263

Db 322 TGGCTAGGAGATCCATGACTTCTCGTGGATGATCTGGAGTACCTCACCTGGGGCATGGAGCA 959  
 /GGGAGGGGCTCTGGTGGCTGAGCTGGGCTGGAGCTGGGAGCTGGGAGA 203

Qy 900 AGGAAGGGAGGCGGCTCTGGGCTTAAGTCAGTGGAGCTGGAGCTGGGAGA 959  
 /GGGAGGGGCTCTGGTGGCTGAGCTGGGCTGGAGCTGGGAGA 203

Db 262 ATGGAGGGGCTCTGGTGGCTGAGCTGGGCTGGAGCTGGGAGCTGGGAGA 203

Qy 960 CGACGCTGGTTAACTTCTCCGCATCCCTCTGGCTGCCCTGGCTGGAGGCTG 1019  
 /GGGGGGTACATGGCCCTGAAGACCCCTCTGGCTCCCATGAGAGCTGCTG 143

Db 202 GTGCGGAGGAGTACGGTGGCG 1042  
 /GTTGCGCCGCTGTCGAGACTGC 120

RESULT 7

CO015154/c

LOCUS CO015154 855 bp mRNA Linear EST 09-JUN-2004  
 DEFINITION EST785336 Coccidioides posadasii saprobic phase cDNA library, greater than 4kb Coccidioides posadasii cDNA clone CIBBG75 5' end, mRNA sequence.

ACCESSION CO015154

VERSION CO015154.1

KEYWORDS EST.

SOURCE Coccidioides posadasii

ORGANISM Coccidioides posadasii

REFERENCE EST785336 Coccidioides posadasii saprobic phase cDNA library, greater than 4kb Coccidioides posadasii cDNA clone CIBBG75 5' end, mRNA sequence.

AUTHORS Gardner,M.J. and Cole,G.T.

TITLE Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags

JOURNAL Unpublished (2003)

COMMENT Other ESTs: EST700694  
 Contact: Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@tigr.org  
 Seq primer: M13 Reverse.

FEATURES source

1. .909  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="taxon:199306"  
 /clone="CIDAK77"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"  
 /note="vector: pEXPRESS 1; Site\_1: Not I; Site\_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated CDNA > 4 kb"

ORIGIN

/dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_lib="Coccidioides posadasii saprobic phase cDNA library, greater than 4kb"  
 /note="vector: pEXPRESS 1; Site\_1: Not I; Site\_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated CDNA > 4 kb"

Query Match 2.3%; Score 47; DB 7; Length 909;

Best Local Similarity 52.9%; Pred. No. 0.66; Mismatches 90; Indels 0; Gaps 0; Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 838 GCGGCTTGTGGAGCATAGGCCATTGGCTTGGAAGGGCTCTGGACATCTGGGGCTATGGA 897  
 Db 884 GTCGCTGGCGCTGTCATGGCCCTGCTGGCTCTGGCCATCTGGGGCTATGGA 825

Qy 898 GAGGACTGTGTTAATCTCTGGCCATTAGGTCTAGGAGCATCTGGGGCTATGGA 897  
 Db 824 GAGCTGTGGCGCTGTCATGGCCCTGCTGGCTCTGGCCATCTGGGGCTATGGA 765

Qy 958 GAGGACTGTGTTAATCTCTGGCCATTAGGTCTAGGAGCATCTGGGGCTATGGA 957  
 Db 764 ATGTCATGGCGCCGCTCTGTGGCCATATGTCATGGCCCTGCTGGGGCT 1017

Qy 1018 GTGTCGCAGC 1028  
 Db 897 GTCCTGGCGCTGTCATGGCCCTGCTGGCCATCTGGGGCTATGGA 838

Qy 898 GAGGACTGTGTTAATCTCTGGCCATTAGGTCTAGGAGCATCTGGGGCTATGGA 957  
 Db 837 GAGCTGTGGCGCTCTGTGGCCATATGTCATGGCCCTGCTGGGGCTATGGA 778

**RESULT 9**  
**CP82948/C**  
**LOCUS** CFP26948 914 bp mRNA linear EST 01-APR-2004  
**DEFINITION** 4 kb Coccidioides posadasii saprobic phase cDNA library, 2 to sequence.  
**ACCESSION** CFP26948  
**VERSION** CFP26948.1 GI:45933005  
**KEYWORDS** EST.  
**SOURCE** Coccidioides posadasii  
**ORGANISM** Coccidioides posadasii  
**COMMENT** Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Onychomycetes; mitosporic Onygenales; Coccidioides.  
**REFERENCE** Gardner, M.J. and Cole, G.T. Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags  
**AUTHORS** Gardner, M.J. and Cole, G.T.  
**TITLE** Analysis of gene expression in Coccidioides posadasii mycelia and spherules via expressed sequence tags  
**JOURNAL** Unpublished (2003)  
**COMMENT** Other\_ESTS: EST700759  
**CONTACT** Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: Gardner@tigr.org  
 Seq primer: M13 Reverse.  
**FEATURES**  
**source**  
 1. .923  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="ITaxon:199306"  
 /clone="CDAL17"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_lib="Coccidioides posadasii saprobic phase cDNA library, 2 to 4 kb"  
 /note="Vector: pExpress 1; Site\_1: Not I; Site\_2: Eco RV; Coccidioides posadasii saprobic phase cDNA library, size fractionated cDNA 2 to 4 kb"  
**ORIGIN**  
 Query Match 2.3%; Score 47; DB 7; Length 923;  
 Best Local Similarity 52.9%; Pred. No. 0.66; Mismatches 90; Indels 0; Gaps 0; Matches 101; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 838 GCGGCTTGTGGAGCATAGGCCATTGGCTTGGAAGGGCTCTGGACATCTGGGGCTATGGA 897  
 Db 907 GAGCTGTGGCGCTGTCATGGCCCTGCTGGCCATCTGGGGCTATGGA 848

Qy 898 GAGGACTGTGTTAATCTCTGGCCATTAGGTCTAGGAGCATCTGGGGCTATGGA 957  
 Db 847 GAGCTGTGGCGCTGTCATGGCCCTGCTGGGGCTATGGA 788

Qy 958 GAGGACTGTGTTAATCTCTGGCCATTAGGTCTAGGAGCATCTGGGGCTATGGA 1017  
 Db 787 ATGTCATGGCGCCGCTGCTGGCCATCTGGGGCTATGGA 728

Qy 1018 GTGTCGCAGC 1028  
 Db 727 GTGTCATGGC 717





The Plant Genome Initiative at Rutgers, Waksman Institute, Rutgers University  
 190 Frelinghuysen Road, Piscataway, NJ 08854, USA  
 Tel: 732 445 3801  
 Fax: 732 445 5735  
 Email: bhari@waksman.rutgers.edu

Seq primer: SP6  
 Class: BAC ends

FEATURES  
 SOURCE  
 High quality sequence start: 394.  
 Location/Qualifiers  
 1. . 753  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /cultivar="B73"  
 /db\_xref="taxon:4577"  
 /clone="ZMBBC0257G23"  
 /lab\_host="E. coli DH1B"  
 /clone\_lib="ZMBBC (EcoRI)"  
 /note="vector: PTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI"

## ORIGIN

Query Match 2.2%; Score 45.4; DB 9; Length 753;  
 Best Local Similarity 54.0%; Pred. No. 1.7;  
 Matches 116; Conservative 0; Mismatches 96; Indels 3; Gaps 1;  
 Qy 775 GTAGCCGCCACTCCTCCCCAGTCCTCACTTGTAGGGCGGGATGCT 834  
 Db 617 GTCGCCGCCAGCTCTGCCTACTGTCCTCGCTGCTGCCCTGGCCATCGTC 558  
 Qy 835 GGTGC---GGTGTGTCAGATAAGCCTGGAAAGGTGCTGTGGACATCTGGGGC 891  
 Db 557 GTTCTCTTCGGCTATCGCATGCCCTCTGCTGTCGTCGACATGGCTGTC 498  
 Qy 892 TATGGAGCAGAGTSCAGGGCGCTGCTGCCCTTAAGGTCATGAGGGGAAATGCC 951  
 Db 497 ACTGTGCCCTCATGCCGCCCTGTATGGCTGCCCTCATGGAGTGGAGGTCC 438  
 Qy 952 TCCACCGAGGAGCTGCTTAACTTACTCCCTCCAT 986  
 Db 437 TCCACCTACATAAGAGCTGGCTCCCTGGAGTCAT 403

Search completed: February 25, 2005, 07:35:09  
 Job time : 6010 secs